



UNIVERSITY OF LEEDS

This is a repository copy of *GSF-locality is not sufficient for proximity-oblivious testing*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/175354/>

Version: Accepted Version

---

**Proceedings Paper:**

Adler, I [orcid.org/0000-0002-9667-9841](https://orcid.org/0000-0002-9667-9841), Köhler, N and Peng, P (Accepted: 2021) GSF-locality is not sufficient for proximity-oblivious testing. In: LIPIcs : Leibniz International Proceedings in Informatics. Computational Complexity Conference (CCC), 20-22 Jul 2021, Online. Schloss Dagstuhl - Leibniz-Zentrum für Informatik . (In Press)

---

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# GSF-locality is not sufficient for proximity-oblivious testing

Isolde Adler ✉

School of Computing, University of Leeds, UK

Noleen Köhler ✉

School of Computing, University of Leeds, UK

Pan Peng ✉ 

Department of Computer Science, University of Sheffield, UK

---

## Abstract

In Property Testing, *proximity-oblivious testers* (POTs) form a class of particularly simple testing algorithms, where a basic test is performed a number of times that may depend on the proximity parameter, but the basic test itself is independent of the proximity parameter.

In their seminal work, Goldreich and Ron [STOC 2009; SICOMP 2011] show that the graph properties that allow constant-query proximity-oblivious testing in the bounded-degree model are precisely the properties that can be expressed as a *generalised subgraph freeness* (GSF) property that satisfies the *non-propagation* condition. It is left open whether the non-propagation condition is necessary. Indeed, calling properties expressible as a generalised subgraph freeness property *GSF-local properties*, they ask whether all GSF-local properties are non-propagating. We give a negative answer by exhibiting a property of graphs that is GSF-local and propagating. Hence in particular, our property does not admit a POT, despite being GSF-local. We prove our result by exploiting a recent work of the authors which constructed a first-order (FO) property that is not testable [SODA 2021], and a new connection between FO properties and GSF-local properties via neighbourhood profiles.

**2012 ACM Subject Classification** Theory of computation → Streaming, sublinear and near linear time algorithms; Theory of computation → Finite Model Theory

**Keywords and phrases** Property testing, proximity-oblivious testing, locality, first-order logic, lower bound

**Digital Object Identifier** 10.4230/LIPIcs.CCC.2021.34

## 1 Introduction

Graph property testing is a framework for studying sampling-based graph algorithms. Given a graph property  $\mathcal{P}$ , the goal is to design a (randomised) algorithm, called *tester*, that distinguishes between graphs that satisfy  $\mathcal{P}$  from those that are ‘far’ from satisfying  $\mathcal{P}$ , where the notion ‘being far’ depends on the underlying query access model and is always parametrised by a *proximity parameter*  $\varepsilon > 0$ . The query model also specifies the class of graphs and the types of queries allowed by the algorithm. The two most well known models for graph property testing are the *dense graph model* and the *bounded-degree graph model* (see [9]). Towards an understanding of which graph properties are testable with a constant number of queries in each model, much progress has been made since the framework of property testing was introduced [24, 10]. To illustrate, a full characterization of the properties that are testable with a constant number of queries in the dense graph model has been obtained by Alon, Fischer, Newman, and Shapira [2].

Typical property testers make decisions regarding the global property of the graph from the local views. In the extreme case, a tester could make local views independent of the distance to a predetermined set of graphs. Motivated by this, Goldreich and Ron [13] initiated

the study of (one-sided error) *proximity-oblivious testers* (POTs) for graphs, where a tester simply repeats a basic test for a number of times that depends on the proximity parameter, and the basic tester is oblivious of the proximity parameter. They gave characterizations of graph properties that can be tested with constant query complexity by a POT in both dense graph model and the bounded-degree model. In each model, it is known that the class of properties that have constant-query POTs is a strict subset of the class of properties that are testable (by standard testers).

In this paper, we focus on the bounded-degree graph model [12]. In this model, the algorithm is given query access to an input graph with maximum degree bounded by  $d$ , where  $d$  is some constant. For any specified query  $v$  and an index  $i \leq d$ , the algorithm can obtain the  $i$ -th neighbor of  $v$  if it exists, and a special symbol  $\perp$  otherwise. Given a proximity parameter  $\varepsilon > 0$ , an  $n$ -vertex graph with maximum degree at most  $d$  is said to be  $\varepsilon$ -far from a property  $\mathcal{P}$  if one needs to add and/or delete more than  $\varepsilon dn$  edges to make it satisfy  $\mathcal{P}$ . A property is said to be *testable* if there exists a tester that makes only a *constant* number of queries to the input graph  $G$ , and distinguishes if  $G$  satisfies the property  $\mathcal{P}$  or is  $\varepsilon$ -far from satisfying  $\mathcal{P}$ , with success probability at least  $\frac{2}{3}$ . Here the constant is a number that might depend on  $\varepsilon$  and  $d$ , but is independent of the size of the input graph. It has been known that many properties are testable, such as subgraph-freeness,  $k$ -edge connectivity, cycle-freeness, being Eulerian, degree-regularity [12], minor-freeness [3, 16, 20], hyperfinite properties [22],  $k$ -vertex connectivity [25, 7], and subdivision-freeness [19].

Turning to POTs, informally, a (one-sided error) POT for a property  $\mathcal{P}$  is a tester that always accepts a graph  $G$  if it satisfies  $\mathcal{P}$ , and rejects  $G$  with probability that is a monotonically increasing function of the distance of  $G$  from the property  $\mathcal{P}$ . We say  $\mathcal{P}$  is *proximity-oblivious testable* if such a tester exists for  $\mathcal{P}$  with constant query complexity. To characterise the class of proximity-oblivious testable properties in the bounded-degree model, Goldreich and Ron [13] introduced a notion of generalized subgraph freeness (GSF), that extends the notions of induced subgraph freeness and (non-induced) subgraph freeness. A graph property is called a *GSF-local* property if it is expressible as a GSF property. It has been shown in [13] that a graph property is constant-query proximity-oblivious testable if and only if it is a GSF-local property that satisfies a so-called *non-propagation* condition. Informally, a GSF-local property  $\mathcal{P}$  is non-propagating if repairing a graph  $G$  that does not satisfy  $\mathcal{P}$  does not trigger a global “chain reaction” of necessary modifications. We refer Section 2.3 for formal definitions.

A major question that is left open is whether every GSF-local property satisfies the non-propagation condition.

## 1.1 Our contribution

In this paper, we resolve the aforementioned open question raised in [13] by showing the following negative result.

► **Theorem 1 (Main result).** *There exists a GSF-local property that is not testable in the bounded-degree graph model. Thus, not all GSF-local properties are non-propagating.*

We expect our result would shed some light on a full characterization of testable properties in the bounded degree model. Indeed, in the recent work by Ito, Khoury and Newman [18], the authors gave a characterization of testable *monotone* graph properties and testable *hereditary* graph properties with one-sided error in the bounded-degree graph model; and they asked the open question “*is every property that is defined by a set of forbidden configurations testable?*”

Since their definition of a property defined by a set of “forbidden configuration” is equivalent to a GSF-local property, our main result also gives a negative answer to their question.

## 1.2 Proof outline

The starting point of our proof is a recent result of the authors that there exists a first-order (FO) property that is not testable in the bounded degree graph model [1], where a property  $\mathcal{P}$  is said to be an FO property if it can be expressed by an FO formula, i. e. a quantified formula whose variables represent graph vertices, with predicates for equality and adjacency. Intuitively, each structure in the property given in [1] is a hybridization of a sequence of expander graphs and a tree structure, where the expander graphs are recursively constructed by the zig-zag product introduced by Reingold et al. [23]. Here each level of the tree structure forms one member of the recursive sequence of expander graphs. It was shown that this property is both an FO property and a family of expanders, and the latter implies it is not testable (see e. g. [6]). We refer to Section 4 and [1] for a detailed description of the property.

By Gaifman’s locality theorem [8], it is known that FO can only express local properties. Indeed, Hanf’s Theorem [15] implies that we can understand this locality as prescribing upper and lower bounds for the occurrence of certain local neighbourhood (isomorphism) types.

On the other hand, a GSF-local property as defined in [13] refers to the freeness of some constant-size *marked* graphs, where a mark graph  $F$  specifies an induced subgraph and how it ‘interacts’ with the rest of the graph (see Definition 3). Intuitively, such a property just specifies a condition that the local neighbourhoods of a graph  $G$  should satisfy, i. e., certain types of local neighbourhoods cannot not occur in  $G$ , or equivalently, these types have 0 occurrences.

Building upon the above observations, we establish a formal connection between FO properties and GSF-local properties. We first encode the possible bounds on occurrences of local neighbourhood types into what we call *neighbourhood profiles*, and characterise FO definable properties of bounded degree relational structures as finite unions of properties defined by neighbourhood profiles (Lemma 9). We then show that every FO formula defined by a non-trivial finite union of properties which in turn is defined by a so-called *0-profiles*, i. e. the prescribed lower bounds are all 0, is GSF-local (Theorem 11). Given the fundamental roles of local properties in graph theory, graph limits [21], we believe this new connection is of independent interest.

For technical reasons, we make use of a property  $\mathcal{P}_{\mathbb{Z}}$  of *relational structures* that can be expressed by some FO formula while it is *not* testable in the bounded-degree model, instead of directly using the non-testable graph property from [1]. We further prove that a minor variant of the relational structure property  $\mathcal{P}_{\mathbb{Z}}$ , which we denote by  $\mathcal{P}'_{\mathbb{Z}}$ , can be defined by 0-profiles (Lemma 20). Finally, we construct a non-testable *graph* property  $\mathcal{P}_{\text{graph}}$  by a *local* reduction from the  $\sigma$ -structure property  $\mathcal{P}'_{\mathbb{Z}}$  (Lemma 24). In the reduction we maintain being definable by 0-profiles which proves GSF-locality of the graph property  $\mathcal{P}_{\text{graph}}$  (Lemma 25). Intuitively, the property  $\mathcal{P}_{\text{graph}}$  encodes the property  $\mathcal{P}_{\mathbb{Z}}$  in undirected graphs. Again,  $\mathcal{P}_{\text{graph}}$  is a family of expanders (which guarantees non-testability), where in addition the local neighbourhoods satisfy the aforementioned features which guarantee that it is an FO property and also GSF-local.

### 1.3 Other related work

The notion of POT was implicitly defined in [4]. Goldreich and Shinkar [14] studied two-sided error POTs for both dense graph and bounded-degree graph models. Goldreich and Kaufman [11] investigated the relation between local conditions that are invariant in an adequate sense and properties that have a constant-query proximity-oblivious testers. Fichtenberger et al. [6] showed that every testable property is either finite or contains an infinite hyperfinite subproperty.

## 2 Preliminaries

### 2.1 Graphs, relational structures and first-order logic

We will briefly introduce structures and first-order logic and point the reader to [5] for a more detailed introduction. A (relational) *signature* is a finite set  $\sigma = \{R_1, \dots, R_\ell\}$  of relation symbols  $R_i$ . Every relation symbol  $R_i$  has an arity  $\text{ar}(R_i) \in \mathbb{N}_{>0}$ . A  $\sigma$ -*structure* is a tuple  $A = (U(A), R_1(A), \dots, R_\ell(A))$ , where  $U(A)$  is a *finite* set, called the *universe* of  $A$  and  $R_i(A) \subseteq U(A)^{\text{ar}(R_i)}$  is an  $\text{ar}(R_i)$ -ary relation on  $U(A)$ . Note that if  $\sigma = \{E_1, \dots, E_\ell\}$  is a signature where each  $E_i$  is a binary relation symbol, then  $\sigma$ -structures are directed graphs with  $\ell$  edge-colours. Let  $\sigma_{\text{graph}} := \{E\}$  be a signature with one binary relation symbol  $E$ . Then we can understand undirected graphs as  $\sigma_{\text{graph}}$ -structures for which the relation  $E$  is symmetric (every undirected edge is represented by two tuples). Using this we can transfer all notions defined below for graphs. Typically we name graphs  $G, H, F$ , we denote the set of vertices of a graph  $G$  by  $V(G)$ , the set of edges by  $E(G)$  and vertices are typically named  $u, v, w, u', v', w', \dots$ . In contrast when we talk about a general relational structure we use  $A, B$  and  $a, b, a', b', \dots$  to denote elements from the universe.

In the following we let  $\sigma$  be a relational signature. Two  $\sigma$ -structures  $A$  and  $B$  are *isomorphic* if there is a bijective map from  $U(A)$  to  $U(B)$  that preserves all relations. For a  $\sigma$ -structure  $A$  and a subset  $S \subseteq U(A)$ , we let  $A[S]$  denote the *substructure* of  $A$  induced by  $S$ , i.e.  $A[S]$  has universe  $S$  and  $R(A[S]) := R(A) \cap S^{\text{ar}(R)}$  for all  $R \in \sigma$ . The *degree* of an element  $a \in U(A)$  denoted by  $\deg_A(a)$  is defined to be the number of tuples in  $A$  containing  $a$ . We define the *degree* of  $A$ , denoted by  $\deg(A)$ , to be the maximum degree of its elements. Given a signature  $\sigma$  and a constant  $d$ , we let  $\mathcal{C}_{\sigma,d}$  be the class of bounded-degree  $d$   $\sigma$ -structures and  $\mathcal{C}_d$  the set of all bounded-degree  $d$  graphs. Note that the degree of a graph differs by exactly a factor 2 from the degree of the corresponding  $\sigma_{\text{graph}}$ -structure.

Syntax and semantic of FO is defined in the usual way (see e.g. [5]). We use  $\exists^{\geq m} x \varphi$  (and  $\exists^{=m} x \varphi$ ,  $\exists^{\leq m} x \varphi$ , respectively) as a shortcut for the FO formula expressing that the number of witnesses  $x$  satisfying  $\varphi$  is at least  $m$  (exactly  $m$ , at most  $m$ , respectively). We say that a variable occurs *freely* in an FO formula if at least one of its occurrences is not bound by any quantifier. We use  $\varphi(x_1, \dots, x_k)$  to express that the set of variables which occur freely in the FO formula  $\varphi$  is a subset of  $\{x_1, \dots, x_k\}$ . For a formula  $\varphi(x_1, \dots, x_k)$ , a  $\sigma$ -structure  $A$  and  $a_1, \dots, a_k \in U(A)$  we write  $A \models \varphi(a_1, \dots, a_k)$  if  $\varphi$  evaluates to true after assigning  $a_i$  to  $x_i$ , for  $1 \leq i \leq k$ . A *sentence* of FO is a formula with no free variables. For an FO sentence  $\varphi$  we say that  $A$  is a *model* of  $\varphi$  or  $A$  satisfies  $\varphi$  if  $A \models \varphi$ .

The *Gaifman graph* of a  $\sigma$ -structure  $A$  is the undirected graph  $G(A) = (U(A), E)$ , where  $\{v, w\} \in E$ , if  $v \neq w$  and there is an  $R \in \sigma$  and a tuple  $\bar{a} = (a_1, \dots, a_{\text{ar}(R)}) \in R(A)$ , such that  $v = a_j$  and  $w = a_k$  for some  $1 \leq j, k \leq \text{ar}(R)$ . We use  $G(A)$  to apply graph theoretic notions to relational structures. Note that for any graph the Gaifman graph of the corresponding symmetric  $\sigma_{\text{graph}}$ -structure is the graph itself. For two elements  $a, b \in U(A)$ , we define the

*distance* between  $a$  and  $b$  in  $A$ , denoted by  $\text{dist}_A(a, b)$ , as the length of a shortest path from  $a$  to  $b$  in  $G(A)$ , or  $\infty$  if there is no such path. For  $r \in \mathbb{N}$  and  $a \in U(A)$ , the  $r$ -neighbourhood of  $a$  is the set  $N_r^A(a) := \{b \in U(A) : \text{dist}_A(a, b) \leq r\}$ . We define  $\mathcal{N}_r^A(a) := A[N_r^A(a)]$  to be the substructure of  $A$  induced by the  $r$ -neighbourhood of  $a$ . For  $r \in \mathbb{N}$  an  $r$ -ball is a tuple  $(B, b)$ , where  $B$  is a  $\sigma$ -structure,  $b \in U(B)$  and  $U(B) = N_r^B(b)$ , i.e.  $B$  has radius  $r$  and  $b$  is the centre. Note that by definition  $(\mathcal{N}_r^A(a), a)$  is an  $r$ -ball for any  $\sigma$ -structure  $A$  and  $a \in U(A)$ . Two  $r$ -balls  $(B, b), (B', b')$  are isomorphic if there is an isomorphism of  $\sigma$ -structure from  $B$  to  $B'$  that maps  $b$  to  $b'$ . We call the isomorphism classes of  $r$ -balls  $r$ -types. For an  $r$ -type  $\tau$  and an element  $a \in U(A)$  we say that  $a$  has  $(r)$ -type  $\tau$  if  $(\mathcal{N}_r^A(a), a) \in \tau$ . Moreover, given such an  $r$ -type  $\tau$ , there is a formula  $\varphi_\tau(x)$  such that for every  $\sigma$ -structure  $A$  and for every  $a \in U(A)$ ,  $A \models \varphi_\tau(a)$  iff  $(\mathcal{N}_r^A(a), a) \in \tau$ . A *Hanf-sentence* is a sentence of the form  $\exists^{\geq m} x \varphi_\tau(x)$ , for some  $m \in \mathbb{N}_{>0}$ , where  $\tau$  is an  $r$ -type. An FO sentence is in *Hanf normal form*, if it is a Boolean combination<sup>1</sup> of Hanf sentences. Two formulas  $\varphi(x_1, \dots, x_k)$  and  $\psi(x_1, \dots, x_k)$  of signature  $\sigma$  are called  $d$ -equivalent, if they are equivalent on  $\mathcal{C}_{\sigma, d}$ , i.e. for all  $A \in \mathcal{C}_{\sigma, d}$  and  $(a_1, \dots, a_k) \in U(A)^k$  we have  $A \models \varphi(a_1, \dots, a_k)$  iff  $A \models \psi(a_1, \dots, a_k)$ . Hanf's locality theorem for first-order logic [15] implies the following.

► **Theorem 2** (Hanf [15]). *Let  $d \in \mathbb{N}$ . Every sentence of first-order logic is  $d$ -equivalent to a sentence in Hanf normal form.*

## 2.2 Property testing

In the following, we give definitions of two models for property testing - the bounded-degree model for graphs and the bounded-degree model for relational structures. For notational convenience,  $\mathcal{C}$  will either denote a class of graphs of bounded-degree  $d$ , or a class of  $\sigma$ -structures of bounded-degree  $d$  for some signature  $\sigma$  and some  $d \in \mathbb{N}$ . We will further refer to both graphs and  $\sigma$ -structures as structures. A *property*  $\mathcal{P}$  in  $\mathcal{C}$  is a subset of  $\mathcal{C}$  which is closed under isomorphism. We say that a structure  $A$  has property  $\mathcal{P}$  if  $A \in \mathcal{P}$ . For  $\epsilon \in (0, 1)$  we say that a structure  $\mathcal{A}$  on  $n$  vertices/elements is  $\epsilon$ -close to  $\mathcal{P}$  if there is a structure  $A' \in \mathcal{P}$  such that  $A$  and  $A'$  differ in at most  $\epsilon dn$  edges/tuples. We say that  $A \in \mathcal{C}$  is  $\epsilon$ -far from  $\mathcal{P}$  if  $A$  is not  $\epsilon$ -close to  $\mathcal{P}$ .

A property tester accesses a structure via oracle queries. A *query* to a  $\sigma$ -structure  $A$  of bounded-degree  $d$  has the form  $(a, i)$  for an element  $a \in U(A)$ ,  $i \in \{1, \dots, d\}$  and is answered by  $\text{ans}(a, i) := (R, a_1, \dots, a_{\text{ar}(R)})$  where  $(a_1, \dots, a_{\text{ar}(R)})$  is the  $i$ -th tuple containing  $a$  and  $(a_1, \dots, a_{\text{ar}(R)}) \in R(A)$ . A *query* to a graph  $G$  of bounded-degree  $d$  has the form  $(v, i)$  for  $v \in V(G)$ ,  $i \in \{1, \dots, d\}$  and is answered by  $\text{ans}(v, i) := w$  where  $w$  is the  $i$ -th neighbour of  $v$ .

Let  $\mathcal{P}_n$  be the subset of  $\mathcal{P}$  with  $n$  vertices/elements. Thus  $\mathcal{P} = \bigcup_{n \in \mathbb{N}} \mathcal{P}_n$ . We give the formal definitions of standard property testing and proximity-oblivious testing in Appendix A.

## 2.3 Generalised subgraph freeness

Now we present the formal definition of generalised subgraph freeness, GSF-local properties and the notion of non-propagation, which were introduced in [13].

► **Definition 3** (Generalized subgraph freeness (GSF)). *A marked graph is a graph with each vertex marked as either ‘full’ or ‘semifull’ or ‘partial’. An embedding of a marked graph*

<sup>1</sup> By Boolean combination we always mean *finite* Boolean combination.



$F$  into a graph  $G$  is an injective map  $f : V(F) \rightarrow V(G)$  such that for every  $v \in V(F)$  the following three conditions hold.

1. If  $v$  is marked ‘full’, then  $N_1^G(f(v)) = f(N_1^F(v))$ .
2. If  $v$  is marked ‘semifull’, then  $N_1^G(f(v)) \cap f(V(F)) = f(N_1^F(v))$ .
3. If  $v$  is marked ‘partial’, then  $N_1^G(f(v)) \supseteq f(N_1^F(v))$ .

The graph  $G$  is called  $F$ -free if there is no embedding of  $F$  into  $G$ . For a set of marked graphs  $\mathcal{F}$ , a graph  $G$  is called  $\mathcal{F}$ -free if it is  $F$ -free for every  $F \in \mathcal{F}$ .

Based on the above definition of GSF, we can define GSF-local properties.

► **Definition 4** (GSF-local properties). Let  $\mathcal{P} = \cup_{n \in \mathbb{N}} \mathcal{P}_n$  be a graph property where  $\mathcal{P}_n = \{G \in \mathcal{P} \mid |V(G)| = n\}$  and  $\overline{\mathcal{F}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  a sequence of sets of marked graphs.  $\mathcal{P}$  is called  $\overline{\mathcal{F}}$ -local if there exists an integer  $s$  such that for every  $n$  the following conditions hold.

1.  $\mathcal{F}_n$  is a set of marked graphs, each of size at most  $s$ .
2.  $\mathcal{P}_n$  equals the set of  $n$ -vertex graphs that are  $\mathcal{F}_n$ -free.

$\mathcal{P}$  is called GSF-local if there is a sequence  $\overline{\mathcal{F}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  of sets of marked graphs such that  $\mathcal{P}$  is  $\overline{\mathcal{F}}$ -local.

The following notion of non-propagating condition of a sequence of sets of marked graphs was introduced to study constant-query POTs.

► **Definition 5** (Non-propagating). Let  $\overline{\mathcal{F}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  be a sequence of sets of marked graphs.

- For a graph  $G$ , a subset  $B \subset V(G)$  covers  $\mathcal{F}_n$  in  $G$  if for every marked graph  $F \in \mathcal{F}_n$  and every embedding of  $F$  in  $G$ , at least one vertex of  $F$  is mapped to a vertex in  $B$ .
- The sequence  $\overline{\mathcal{F}}$  is non-propagating if there exists a (monotonically non-decreasing) function  $\tau : (0, 1] \rightarrow (0, 1]$  such that the following two conditions hold.

1. For every  $\epsilon > 0$  there exists  $\beta > 0$  such that  $\tau(\beta) < \epsilon$ .
2. For every graph  $G$  and every  $B \subset V(G)$  such that  $B$  covers  $\mathcal{F}_n$  in  $G$ , either  $G$  is  $\tau(|B|/n)$ -close to being  $\mathcal{F}_n$ -free or there are no  $n$ -vertex graphs that are  $\mathcal{F}_n$ -free.

A GSF-local property  $\mathcal{P}$  is non-propagating if there exists a non-propagating sequence  $\overline{\mathcal{F}}$  such that  $\mathcal{P}$  is  $\overline{\mathcal{F}}$ -local.

In the above definition, the set  $B$  can be viewed as the set involving necessary modifications for repairing a graph  $G$  that does not satisfy the property  $\mathcal{P}$  that is  $\overline{\mathcal{F}}$ -local, and the second condition says we do not need to modify  $G$  “much beyond”  $B$ . In particular, it implies we can repair  $G$  without triggering a global “chain reaction”. Goldreich and Ron gave the following characterization for the proximity-oblivious testable properties in the bounded-degree graph model.

► **Theorem 6** (Theorem 5.5 in [13]). A graph property  $\mathcal{P}$  has a constant-query proximity-oblivious tester if and only if  $\mathcal{P}$  is GSF-local and non-propagating.

The following open question was raised in [13].

► **Open Question 7** (Are all GSF-local properties non-propagating?). Is it the case that for every GSF-local property  $\mathcal{P} = \cup_{n \in \mathbb{N}} \mathcal{P}_n$ , there is a sequence  $\overline{\mathcal{F}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  that is non-propagating and  $\mathcal{P}$  is  $\overline{\mathcal{F}}$ -local?

### 3 Relating different notions of locality

In this section we define properties by prescribing upper and lower bounds on the number of occurrence of neighbourhood types. These bounds are given by *neighbourhood profiles* which we will define formally below. We use these properties to give a natural characterization of FO properties of bounded-degree structures in Lemma 9, which is a straightforward consequence of Hanf's Theorem (Theorem 2). We use this characterization to establish links between FO definability and GSF-locality. This connection is the key ingredient in the proof of our main theorem.

Observe that for fixed  $r, d \in \mathbb{N}$  and  $\sigma$ , there are only finitely many  $r$ -types in structures in  $\mathcal{C}_{\sigma,d}$ . For any signature  $\sigma$  and  $d, r \in \mathbb{N}$  we let  $n_{d,r,\sigma} \in \mathbb{N}$  be the number of different  $r$ -types of  $\sigma$ -structures of degree at most  $d$ . Assuming that for all  $d, r \in \mathbb{N}$  the  $r$ -neighbourhood-types of  $\sigma$ -structures of degree at most  $d$  are ordered, we let  $\tau_{d,r,\sigma}^i$  denote the  $i$ -th such neighbourhood type, for  $i \in \{1, \dots, n_{d,r,\sigma}\}$ . With each  $\sigma$ -structure  $A \in \mathcal{C}_{\sigma,d}$  we associate its  $r$ -*histrogram vector*  $\bar{v}_{d,r,\sigma}(A)$ , given by

$$(\bar{v}_{d,r,\sigma}(A))_i := |\{a \in U(A) \mid \mathcal{N}_r^A(a) \in \tau_{d,r,\sigma}^i\}|.$$

We let

$$\mathfrak{I} := \{[k, l], [k, \infty) \mid k \leq l \in \mathbb{N}\}$$

be the set of all closed or half-closed, infinite intervals with natural lower/upper bounds.

► **Definition 8.** Let  $\sigma$  be a signature and  $d, r \in \mathbb{N}$ .

1. An  $r$ -neighbourhood profile of degree  $d$  is a function  $\rho : \{1, \dots, n_{d,r,\sigma}\} \rightarrow \mathfrak{I}$ .
2. For a structure  $A \in \mathcal{C}_{\sigma,d}$ , we say  $A$  obeys  $\rho$ , denoted by  $A \sim \rho$ , if

$$(\bar{v}_{d,r,\sigma}(A))_i \in \rho(i) \text{ for all } i \in \{1, \dots, n_{d,r,\sigma}\}.$$

Let  $\mathcal{P}_\rho$  be the set of structures  $A$  that obey  $\rho$ , i.e.,  $\mathcal{P}_\rho = \{A \in \mathcal{C}_{\sigma,d} \mid A \sim \rho\}$ .

3. We say that a property  $\mathcal{P}$  is defined by a finite union of neighbourhood profiles if there is  $k \in \mathbb{N}$  such that  $\mathcal{P} = \bigcup_{1 \leq i \leq k} \mathcal{P}_{\rho_i}$  where  $\rho_i$  is an  $r_i$ -neighbourhood profile and  $r_i \in \mathbb{N}$  for every  $i \in \{1, \dots, k\}$ .

We let  $n_{d,r} := n_{d,r,\sigma_{\text{graph}}}$  denote the total number of  $r$ -type of undirected graphs of degree at most  $d$ , and let  $\tau_{d,r}^i := \tau_{d,r,\sigma_{\text{graph}}}^i$  be the  $i$ -th  $r$ -type of bounded degree  $d$ , for any  $i \in \{1, \dots, n_{d,r}\}$ . Further, for a graph  $G$  let  $\bar{v}_{d,r}(G)$  denote the  $r$ -histrogram vector of  $G$ . Note that for any type  $\tau_{d,r}^i$  where the edge relation is not symmetric we have that  $(\bar{v}_{d,r}(G))_i = 0$  and therefore in any  $r$ -neighbourhood profile  $\rho$  for graphs we have  $\rho(i) = [0, 0]$  for any type  $\tau_{d,r}^i$  which is not symmetric.

We now give a lemma showing that bounded-degree FO properties can be equivalently defined as finite unions of properties defined by neighbourhood profiles. Here the technicalities that arise are due to Hanf normal form not requiring the locality-radius of all Hanf-sentences to be the same. The proof of Lemma 9 is deferred to Appendix C.

► **Lemma 9.** For every non-empty property  $\mathcal{P} \subseteq \mathcal{C}_{\sigma,d}$ ,  $\mathcal{P}$  is FO definable on  $\mathcal{C}_{\sigma,d}$  if and only if  $\mathcal{P}$  can be obtained as a finite union of properties defined by neighbourhood profiles.



### 3.1 Relating FO properties to GSF-local properties

We now prove that FO properties which arise as unions of neighbourhood profiles of a particularly simple form are GSF-local. For this let

$$\mathcal{I}_0 := \{[0, \infty), [0, k] \mid k \in \mathbb{N}\} \subset \mathcal{I}.$$

We call any neighbourhood profile  $\rho$  with codomain  $\mathcal{I}_0$  a *0-profile*, as all lower bounds for the occurrence of types are 0.

► **Observation 10.** *Let  $\rho$  be a 0-profile. If two structures  $A, A' \in \mathcal{C}_{\sigma, d}$  satisfy  $(\bar{v}_{d, r, \sigma}(A))_i \leq (\bar{v}_{d, r, \sigma}(A'))_i$  for every  $i \in \{1, \dots, n_{d, r, \sigma}\}$  and  $A' \sim \rho$ , then  $A \sim \rho$ .*

*In particular, the existence of an  $r$ -type cannot be expressed by a 0-profile.*

► **Theorem 11.** *Every finite union of properties defined by 0-profiles is GSF-local.*

**Proof.** We prove this in two parts (Claim 12 and Claim 13). We first argue that every property  $\mathcal{P}_\rho$  defined by some 0-profile  $\rho : \{1, \dots, n_{d, r, \sigma}\} \rightarrow \mathcal{I}_0$  is GSF-local. For this it is important to note that we can express a forbidden  $r$ -type  $\tau$  by a forbidden generalised subgraph. For  $(B, b) \in \tau$ , the set of all graphs with no vertex of neighbourhood type  $\tau$  is the set of all  $B$ -free graphs where every vertex in  $V(B)$  of distance less than  $r$  to  $b$  is marked ‘full’ and every vertex in  $V(B)$  of distance  $r$  to  $b$  is marked ‘semifull’. Since a profile of the form  $\rho : \{1, \dots, n_{d, r, \sigma}\} \rightarrow \mathcal{I}_0$  can express that some neighbourhood type  $\tau$  can appear at most  $k$  times for some fixed  $k \in \mathbb{N}$ , we need to forbid all marked graphs in which type  $\tau$  appears  $k + 1$  times. We will formalise this in the following claim.

► **Claim 12.** For every  $r$ -neighbourhood profile  $\rho : \{1, \dots, n_{d, r}\} \rightarrow \mathcal{I}_0$ , there is a finite set  $\mathcal{F}$  of marked graphs such that  $\mathcal{P}_\rho$  is exactly the property of  $\mathcal{F}$ -free graphs.

**Proof.** Assume  $\tau$  is an  $r$ -type and  $k \in \mathbb{N}_{>0}$ . Then we say that a marked graph  $F$  is a  *$k$ -realisation* of  $\tau$  if  $F$  has the following properties.

1. There are  $k$  distinct vertices  $v_1, \dots, v_k$  in  $F$  such that  $(\mathcal{N}_r^F(v_i), v_i) \in \tau$  for every  $i = 1, \dots, k$ .
2. Every vertex  $v$  in  $F$  has distance less or equal to  $r$  to at least one vertex  $v_i$ .
3. Every vertex  $v$  in  $F$  of distance less than  $r$  to at least one  $v_i$  is marked as ‘full’.
4. Every vertex  $v$  in  $F$  of distance greater or equal to  $r$  to every  $v_i$  is marked as ‘semifull’.

We denote by  $S^k(\tau)$  the set of all  $k$ -realisations of  $\tau$ .

Now we can define the set  $\mathcal{F}$  of forbidden subgraphs to be

$$\mathcal{F} := \bigcup_{k \in \mathbb{N}, 1 \leq i \leq n_{d, r, \sigma} : \rho(i) = [0, k]} S^{k+1}(\tau_{d, r}^i).$$

Let  $\mathcal{P}$  be the property of all  $\mathcal{F}$ -free graphs. We first prove that the property  $\mathcal{P}$  is contained in  $\mathcal{P}_\rho$ . Towards a contradiction assume that  $G \in \mathcal{C}_d$  is  $\mathcal{F}$ -free but not contained in  $\mathcal{P}_\rho$ . As  $G$  is not contained in  $\mathcal{P}_\rho$  there must be an index  $i \in \{1, \dots, n_{d, r}\}$  such that  $(\bar{v}_{d, r}(G))_i \notin \rho(i)$ . Since  $\rho(i) \in \mathcal{I}_0$  there is  $k \in \mathbb{N}$  such that  $\rho(i) = [0, k]$  and hence  $(\bar{v}_{d, r}(G))_i > k$ . Hence there must be  $k + 1$  vertices  $v_1, \dots, v_{k+1}$  in  $G$  such that  $(\mathcal{N}_r^G(v_i), v_i) \in \tau_{d, r}^i$ . We define the marked graph  $F$  to be the subgraph of  $G$  induced by the  $r$ -neighbourhoods of  $v_1, \dots, v_{k+1}$ , i.e.  $G[\bigcup_{1 \leq i \leq k+1} \mathcal{N}_r^G(v_i)]$ , in which every vertex of distance less than  $k$  to at least one of the  $v_i$  is marked as ‘full’ and every other vertex is marked as ‘semifull’. Then  $F$  is by definition a  $(k + 1)$ -realisation of  $\tau_{d, r}^i$  and hence  $F \in \mathcal{F}$ . We now argue that  $F$  can be embedded into  $G$ . Since  $F$  is an induced subgraph of  $G$  the identity map gives us a natural

embedding  $f : F \rightarrow G$ . Let  $v$  be any vertex marked ‘full’ in  $F$ . Then by construction of  $F$ , there is  $i \in \{1, \dots, k+1\}$  such that  $f(v)$  is of distance less than  $r$  to  $v_i$  in  $G$ . But then  $N_1^G(f(v))$  is a subset of  $N_r^G(v_i)$ . As  $F$  without the marking is the subgraph of  $G$  induced by  $\cup_{1 \leq i \leq k+1} N_r^G(v_i)$  this implies that  $f(N_1^F(v)) = N_1^G(f(v))$ . Furthermore, assume  $v$  is a vertex marked ‘semifull’ in  $F$ . Then  $f(N_1^F(v)) = N_1^G(f(v)) \cap f(V(F))$  holds as  $F$  without the markings is an induced subgraph of  $G$ . This proves that  $G$  is not  $F$ -free by Definition 3. This is a contradiction to our assumption that  $G$  is  $\mathcal{F}$ -free and  $F \in \mathcal{F}$ .

Similarly, we can show that  $\mathcal{P}_\rho \subseteq \mathcal{P}$  by assuming  $G \in \mathcal{C}_d$  is in  $\mathcal{P}_\rho$  but not  $\mathcal{F}$ -free, and showing that the embedding of any graph of  $\mathcal{F}$  into  $G$  yields an amount of vertices of a certain type contradicting containment in  $\mathcal{P}_\rho$ .  $\triangleleft$

Next we prove that classes defined by excluding finitely many marked graphs are closed under finite unions.

$\triangleright$  **Claim 13.** Let  $\mathcal{F}_1, \mathcal{F}_2$  be two finite sets of marked graphs. For  $i \in \{1, 2\}$ , let  $\mathcal{P}_i$  be the property of  $\mathcal{F}_i$ -free graphs. Then there is a set  $\mathcal{F}$  of generalised subgraphs such that  $\mathcal{P}_1 \cup \mathcal{P}_2$  is the property of  $\mathcal{F}$ -free graphs.

*Proof.* We say that a marked graph  $F$  is a (not necessarily disjoint) union of marked graphs  $F_1, F_2$  if

1. there is an embedding  $f_i$  of  $F_i$  into the graph  $F$  without its markings as in Definition 3 for every  $i \in \{1, 2\}$ .
2. for every vertex  $v$  in  $F$  there is  $i \in \{1, 2\}$  and a vertex  $w$  in  $F_i$  such that  $f_i(w) = v$ .
3. every vertex  $v$  in  $F$  is marked ‘full’, if there is  $i \in \{1, 2\}$  and a ‘full’ vertex  $w$  in  $F_i$  such that  $f_i(w) = v$ .
4. every vertex  $v$  in  $F$  is marked ‘semifull’, if there is  $i \in \{1, 2\}$  and a ‘semifull’ vertex  $w$  in  $F_i$  such that  $f_i(w) = v$  and  $f_i(u) \neq v$  for every  $i \in \{1, 2\}$  and every ‘full’ vertex  $u$ .
5. every vertex  $v$  in  $F$  is marked ‘partial’ if  $f_i(u) \neq v$  for every  $i \in \{1, 2\}$  and every ‘full’ or ‘semifull’ vertex  $u$ .

We define  $S(F_1, F_2)$  to be the set of all possible (not necessarily disjoint) unions of  $F_1, F_2$ . We can now define the set  $\mathcal{F}$  to be

$$\mathcal{F} := \bigcup_{F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2} S(F_1, F_2).$$

Let  $\mathcal{P}$  be the property of all  $\mathcal{F}$ -free graphs. Now we prove  $\mathcal{P} \subseteq \mathcal{P}_1 \cup \mathcal{P}_2$ . Towards a contradiction assume  $G$  is  $\mathcal{F}$ -free but  $G$  is in neither  $\mathcal{P}_1$  nor in  $\mathcal{P}_2$ . Then for every  $i \in \{1, 2\}$  there is a graph  $F_i \in \mathcal{F}_i$  such that  $G$  is not  $F_i$ -free. It is easy to see that there is a union  $F_\cup$  of  $F_1$  and  $F_2$  such that  $G$  is not  $F_\cup$ -free, which contradicts that  $G$  is  $\mathcal{F}$ -free.

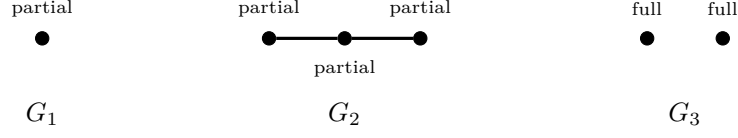
Conversely, in order to prove  $\mathcal{P}_1 \cup \mathcal{P}_2 \subseteq \mathcal{P}$ , if  $G$  is  $\mathcal{F}_i$  free for some  $i \in \{1, 2\}$  then  $G$  must be  $\mathcal{F}$ -free by construction of  $\mathcal{F}$ .  $\triangleleft$

Combining the two claims above proves the Theorem 11.  $\blacktriangleleft$

### Further discussion of the relation between FO and GSF-locality

First let us remark that it is neither true that every FO definable property is GSF-local, nor that every GSF-local property is FO definable.

$\blacktriangleright$  **Example 14.** The property of bounded-degree graphs containing a triangle is FO definable but not GSF-local.



■ **Figure 1** Marked graphs for Example 16.

Indeed, the existence of a fixed number of vertices of certain neighbourhood types can be expressed in FO, while in general, this cannot be expressed by forbidding generalised subgraphs. If a formula has a 0-profile (and hence does not require the existence of any types) then the property defined by that formula is GSF-local, as shown in Theorem 11.

► **Example 15.** The class of all bounded-degree graphs with an even number of vertices is GSF-local but not FO definable.

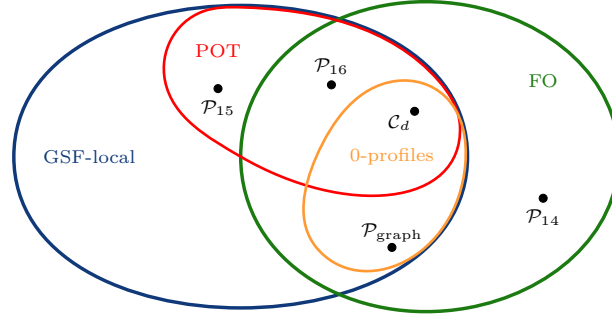
Let us remark that Theorem 11 combined with Lemma 9 proves that every finite union of properties definable by 0-profiles is both FO definable and GSF-local. Hence it is natural to ask whether the intersection of FO definable properties and GSF-local properties is precisely the set of finite unions of properties definable by 0-profiles. However, this is not the case. The following example shows that there are properties which are both FO definable and GSF-local but cannot be expressed by 0-profiles.

► **Example 16.** We let  $d \geq 2$  and let  $B_1 := (\{v\}, \{\})$ ,  $B_2 = (\{v, w\}, \{\{v, w\}\})$  be two graphs. We further let  $\tau_1, \tau_2$  be the 1-types of degree  $d$  such that  $(B_1, v) \in \tau_1$  and  $(B_2, v) \in \tau_2$ . Consider the property  $\mathcal{P}$  defined by the following FO formula

$$\varphi := \neg \exists x(x = x) \vee \exists^{\neq 1} x(\varphi_{\tau_1}(x) \wedge \forall y(x \neq y \rightarrow \varphi_{\tau_2}(y))).$$

$\mathcal{P}$  contains, besides the empty graph, unions of an arbitrary amount of disjoint edges and one isolated vertex. To define a sequence of forbidden subgraphs we let  $G_1, G_2, G_3$  be the marked graphs in Figure 1. Let  $\mathcal{F}_{\text{even}} := \{G_1\}$  and  $\mathcal{F}_{\text{odd}} := \{G_2, G_3\}$  and let  $\overline{\mathcal{F}} = (\mathcal{F}_n)_{n \in \mathbb{N}}$  where  $\mathcal{F}_i = \mathcal{F}_{\text{even}}$  if  $i$  is even and  $\mathcal{F}_i = \mathcal{F}_{\text{odd}}$  if  $i$  is odd. Note that every graph on more than one vertex with an odd number of vertices which is  $\mathcal{F}_{\text{odd}}$ -free must contain a vertex of neighbourhood type  $\tau_1$ , and that the set of  $\mathcal{F}_{\text{even}}$ -free graphs contains only the empty graph. Hence  $\mathcal{P}$  is  $\overline{\mathcal{F}}$ -local. Now assume towards a contradiction that  $\mathcal{P} = \cup_{1 \leq i \leq k} \mathcal{P}_{\rho_i}$  for 0-profiles  $\rho_i$ . Let  $G_m$  be the graph consisting of  $m$  disjoint edges and one isolated vertex and  $H_m$  the graph consisting of  $m$  disjoint edges. Since  $G_m \in \mathcal{P}$  there is  $i \in \{1, \dots, k\}$  such that  $G_m \sim \rho_i$ . By choice of  $G_m$  and  $H_m$  we have  $0 \leq (\overline{v}_{d,r}(H_m))_j \leq (\overline{v}_{d,r}(G_m))_j \in \rho_i(j)$  for every  $j \in \{1, \dots, n_{d,r}\}$ . Since additionally  $\rho_i(j) \in \mathcal{I}_0$  this implies that  $(\overline{v}_{d,r}(H_m))_j \in \rho_i(j)$ . But then  $H_m \sim \rho_i$  which yields a contradiction as  $H_m \notin \mathcal{P}$ . Hence  $\mathcal{P}$  can not be defined as a finite union of 0-profiles.

Figure 2 gives a schematic overview of all classes of properties discussed here and their relationship.



■ **Figure 2** Overview of the classes of properties, here  $\mathcal{P}_i$  refers to the property from Example  $i$ ,  $\mathcal{C}_d$  refers to the property of all graphs of bounded degree  $d$  and  $\mathcal{P}_{\text{graph}}$  is the property defined in Section 4.2.

#### 4 Proof of the main theorem

In this section we prove Theorem 1. We start by describing a property of relational structures, similar to a property in [1], which is not testable. We then show that the property can be expressed by a union of 0-profiles, and hence by Theorem 11 it is GSF-local.

Let  $\sigma$  be the signature,  $d \in \mathbb{N}$  and  $\mathcal{P}_{\mathbb{Z}}$  be the property of  $d$   $\sigma$ -structures of bounded-degree from [1].

##### Brief Description of the property $\mathcal{P}_{\mathbb{Z}}$

$\mathcal{P}_{\mathbb{Z}}$  is the property of all bounded-degree  $d$   $\sigma$ -structures, which satisfy some first-order logic formula  $\varphi_{\mathbb{Z}}$ . On a high level, each structure  $A$  in the property  $\mathcal{P}_{\mathbb{Z}}$  is a hybridization of a sequence of expander graphs and a tree structure, where the expander graphs are constructed by the zig-zag product that was introduced in [23]. Slightly more precisely, each model of  $\varphi_{\mathbb{Z}}$  is a rooted  $k$ -ary complete tree for some constant  $k$ , where the vertices on each level form an expander. In terms of logic language, for some constant  $D > 1$ , we considered

$$\sigma := \{\{E_{i,j}\}_{i,j \in [D]^2}, \{F_k\}_{k \in ([D]^2)^2}, R, \{L_k\}_{k \in ([D]^2)^2}\},$$

where  $E_{i,j}$ ,  $F_k$ ,  $R$  and  $L_k$  are binary relation symbols for  $i, j \in [D]^2$  and  $k \in ([D]^2)^2$ . We further use  $F$  and  $E$  as an abbreviation to denote  $\bigcup_{i,j \in [D]^2} E_{i,j}$  and  $\bigcup_{k \in ([D]^2)^2} F_k$ . We defined an FO formula  $\varphi_{\mathbb{Z}}$  such that

$$\varphi_{\mathbb{Z}} := \varphi_{\text{tree}} \wedge \varphi_{\text{rotationMap}} \wedge \varphi_{\text{base}} \wedge \varphi_{\text{recursion}}, \text{ and } \mathcal{P}_{\mathbb{Z}} := \{\mathcal{A} \in \mathcal{C}_{\sigma,d} \mid \mathcal{A} \models \varphi_{\mathbb{Z}}\},$$

where  $\varphi_{\text{tree}}$ ,  $\varphi_{\text{rotationMap}}$ ,  $\varphi_{\text{base}}$ ,  $\varphi_{\text{recursion}}$  are FO formulas which encode the tree structure (and degree regularity), rotation maps, base graph (with constant size) and recursive construction of expander graphs (via the zig-zag product). Note that for the construction we use some base graph  $H$  which is given by its rotation map  $\text{ROT}_H : ([D]^2)^2 \times [D] \rightarrow ([D]^2)^2 \times [D]$ , which is a special type of an encoding of a graph.

The precise formula is given in Appendix B. We will restate parts of the formula, whenever they are relevant in the proofs below.

#### 4.1 Characterisation by neighbourhood profiles

Our aim in this section is to prove that a minor variation of property  $\mathcal{P}_{\mathbb{Z}}$  of relational structures can be written as a finite union of properties defined by 0-profiles of radius 2. As the existence of a certain vertex cannot be expressed with a 0-profile (see Observation 10) and  $\varphi_{\mathbb{Z}}$  demands the existence of a certain vertex (the root vertex), the property  $\mathcal{P}_{\mathbb{Z}}$  cannot be expressed in terms of 0-profiles. However we define a slight variation of the formula  $\varphi_{\mathbb{Z}}$  which, as we will see later, can be expressed by 0-profiles. Let

$$\varphi'_{\mathbb{Z}} := \varphi'_{\text{tree}} \wedge \varphi_{\text{rotationMap}} \wedge \varphi_{\text{base}} \wedge \varphi_{\text{recursion}},$$

where we obtain  $\varphi'_{\text{tree}}$  from  $\varphi_{\text{tree}}$  by replacing the subformula  $\exists^1 x \varphi_{\text{root}}(x)$  by  $\exists^{\leq 1} x \varphi_{\text{root}}(x)$ , where  $\varphi_{\text{root}}(x) := \forall y \neg F(y, x)$ . We define the property

$$\mathcal{P}'_{\mathbb{Z}} := \{A \in \mathcal{C}_{\sigma, d} \mid A \models \varphi'_{\mathbb{Z}}\}.$$

We denote the empty structure by  $A_{\emptyset}$  (i. e.  $U(A_{\emptyset}) = \emptyset$ ).

► **Lemma 17.** *The properties  $\mathcal{P}'_{\mathbb{Z}}$  and  $\mathcal{P}_{\mathbb{Z}} \cup \{A_{\emptyset}\}$  are equal.*

To prove this we use the following lemma [1, Lemma 3.5].

► **Lemma 18 ([1]).** *For  $A \in \mathcal{C}_{\sigma, d}$  let  $G_F^A$  be the graph with vertex set  $U(A)$  and edge set  $\{(a, b) \mid (a, b) \in F(A)\}$ . If  $A \models \varphi_{\mathbb{Z}}$  then  $G_F^A$  is connected.*

**Proof of Lemma 17.** We first prove that  $\mathcal{P}'_{\mathbb{Z}} \subseteq \mathcal{P}_{\mathbb{Z}} \cup \{A_{\emptyset}\}$ . Consider the formula  $\tilde{\varphi}_{\mathbb{Z}}$  which is obtained from  $\varphi_{\mathbb{Z}}$  by removing the subformula  $\exists^1 x \varphi_{\text{root}}(x)$ . We use the following simple observation, which we will prove in Appendix D.

▷ **Claim 19.** Satisfying  $\tilde{\varphi}_{\mathbb{Z}}$  is closed under disjoint unions on  $\mathcal{C}_{\sigma, d}$ .

Since  $A_{\emptyset} \in \mathcal{P}_{\mathbb{Z}} \cup \{A_{\emptyset}\}$  it is sufficient to consider only non-empty structures in the following. Therefore assume that there exists  $A \in \mathcal{C}_{\sigma, d}$  with  $U(A) \neq \emptyset$  such that  $A \models \varphi'_{\mathbb{Z}}$  and  $A$  contains no element  $a$  for which  $A \models \varphi_{\text{root}}(a)$ . Let  $A' \in \mathcal{C}_{\sigma, d}$  be any model of  $\varphi_{\mathbb{Z}}$  with  $U(A) \cap U(A') = \emptyset$ . Then  $A \cup A' \models \tilde{\varphi}_{\mathbb{Z}}$  by Claim 19. Furthermore,  $A \cup A' \models \exists^1 x \varphi_{\text{root}}(x)$ , which implies  $A \cup A' \models \varphi_{\mathbb{Z}}$ . By construction  $G_F^{A \cup A'}$  has more than one connected component as both  $U(A) \neq \emptyset$  and  $U(A') \neq \emptyset$  and  $A \cup A'$  is a disjoint union of  $A$  and  $A'$ . Hence we obtain a contradiction to Lemma 18. Therefore every non-empty structure satisfying  $\varphi'_{\mathbb{Z}}$  must satisfy  $\exists^1 x \varphi_{\text{root}}(x)$ , and hence also  $\varphi_{\mathbb{Z}}$ .

Conversely, if  $A \in \mathcal{C}_{\sigma, d}$  is a model of  $\varphi_{\mathbb{Z}}$  then  $A \models \exists^1 x \varphi_{\text{root}}(x)$ . This implies directly that  $A \models \exists^{\leq 1} x \varphi_{\text{root}}(x)$  and hence  $A \models \varphi'_{\mathbb{Z}}$ . Furthermore,  $A_{\emptyset} \in \mathcal{P}'_{\mathbb{Z}}$  as  $A_{\emptyset} \models \exists^{\leq 1} x \varphi_{\text{root}}(x)$  and  $A_{\emptyset} \models \tilde{\varphi}_{\mathbb{Z}}$  as  $\tilde{\varphi}_{\mathbb{Z}}$  is a conjunction of universally quantified formulas. Hence  $\mathcal{P}_{\mathbb{Z}} \cup \{A_{\emptyset}\} \subseteq \mathcal{P}'_{\mathbb{Z}}$ . ◀

We now define the 0-profiles which express the property  $\mathcal{P}'_{\mathbb{Z}}$ . For all  $\sigma$ -structures in  $\mathcal{P}_{\mathbb{Z}}$  (all  $\sigma$ -structure in  $\mathcal{P}'_{\mathbb{Z}}$  but  $A_{\emptyset}$ ) it is crucial that they are allowed to contain precisely one root element. Hence the neighbourhood profile describing  $\mathcal{P}'_{\mathbb{Z}}$  must restrict the number of occurrences of the 2-type of the root element. But since in  $\mathcal{P}_{\mathbb{Z}}$ , the root elements in different structures may have different 2-types, we partition  $\mathcal{P}_{\mathbb{Z}}$  into parts  $\mathcal{P}_1, \dots, \mathcal{P}_m$  by the 2-type

of the root element. Note that the number  $m$  of parts is constant as there are at most  $n_{d,2,\sigma}$  2-types in total. For each of these parts we then define a neighbourhood profile  $\rho_k$  such that  $\mathcal{P}_k \cup \{A_\emptyset\} = \mathcal{P}_{\rho_k}$ . We would like to remark here that the roots of all but one structure in  $\mathcal{P}_{\mathbb{Z}}$  actually have the same 2-types. However, proving this requires a detailed insight into the construction of  $\mathcal{P}_{\mathbb{Z}}$ , so we avoid this here and use the partition into finitely many parts instead. We now define the parts and corresponding profiles formally.

Assume without loss of generality that the 2-types  $\tau_{d,2,\sigma}^1, \dots, \tau_{d,2,\sigma}^{n_{d,2,\sigma}}$  of degree  $d$  are ordered in such a way that for  $(B, b) \in \tau_{d,2,\sigma}^k$ , it holds that  $B \models \varphi_{\text{root}}(b)$  if and only if  $k \in \{1, \dots, m\}$  for some  $m \leq n_{d,2,\sigma}$ . For  $k \in \{1, \dots, m\}$ , let

$$\mathcal{P}_k := \{A \in \mathcal{P}_{\mathbb{Z}} \mid \text{there is } a \in U(A) \text{ such that } (\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^k\}.$$

Since every  $A \in \mathcal{P}_{\mathbb{Z}}$  satisfies  $\exists^=1 x \varphi_{\text{root}}(x)$  we get that

$$\mathcal{P}'_{\mathbb{Z}} = \bigcup_{1 \leq k \leq m} \mathcal{P}_k \cup \{A_\emptyset\}$$

and this union is disjoint. Furthermore, for  $k \in \{1, \dots, m\}$ , let  $I_k \subseteq \{1, \dots, n_{d,2,\sigma}\}$  be the set of indices  $j$  such that there is a structure  $A \in \mathcal{P}_k$  and  $a \in U(A)$  with  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^j$ . For every  $k \in \{1, \dots, m\}$  we define the 2-neighbourhood profile  $\rho_k : \{1, \dots, n_{d,2,\sigma}\} \rightarrow \mathcal{I}_0$  by

$$\rho_k(i) := \begin{cases} [0, 1] & \text{if } i = k, \\ [0, \infty) & \text{if } i \in I_k \setminus \{k\}, \\ [0, 0] & \text{otherwise.} \end{cases}$$

To prove that these 0-profiles of radius 2 define the property  $\mathcal{P}'_{\mathbb{Z}}$ , the crucial observation is that for every element  $a$  of some structure in  $\mathcal{C}_{\sigma,d}$ , the FO-formula  $\varphi'_{\mathbb{Z}}$  only talks about elements of distance at most 2 to  $a$  (i.e.  $\varphi'_{\mathbb{Z}}$  is 2-local). Hence the 2-histogram vector of a structure already captures whether the structure satisfies  $\varphi'_{\mathbb{Z}}$ . We will now formally prove this.

► **Lemma 20.** *It holds that  $\mathcal{P}'_{\mathbb{Z}} = \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$ .*

**Proof.** We first prove that  $\mathcal{P}'_{\mathbb{Z}} \subseteq \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$ . First note that trivially  $A_\emptyset \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$ . Now assume  $A \in \mathcal{P}_{\mathbb{Z}}$ . This implies that there is  $k \in \{1, \dots, m\}$  such that  $A \in \mathcal{P}_k$ . By construction we have that for every  $a \in A$ , there is  $i \in I_k$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^i$ . Furthermore, since  $A \models \varphi_{\mathbb{Z}}$ , we have that  $A \models \exists^=1 x \varphi_{\text{root}}(x)$ , and that there can be at most one  $a \in U(A)$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^k$ . Therefore  $A \in \mathcal{P}_{\rho_k}$ .

To prove  $\bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \subseteq \mathcal{P}'_{\mathbb{Z}}$ , we prove that every structure in  $\bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$  must satisfy  $\varphi'_{\mathbb{Z}}$ . We will prove that every  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$  satisfies  $\varphi_{\text{recursion}}$ , and refer for the proof that  $A$  satisfies  $\varphi'_{\text{tree}} \wedge \varphi_{\text{rotationMap}} \wedge \varphi_{\text{base}}$  to Claim 30, Claim 31 and Claim 32 in Appendix D. Note that  $A_\emptyset \models \varphi'_{\mathbb{Z}}$  by Lemma 17 and hence we exclude  $A_\emptyset$  in the following.

► **Claim 21.** Every structure  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$  satisfies  $\varphi_{\text{recursion}}$ .

**Proof.** Let  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$ . Then there is a  $k \in \{1, \dots, m\}$  such that  $A \in \mathcal{P}_{\rho_k}$ .

### 34:14 GSF-locality is not sufficient for proximity-oblivious testing

By definition,  $\varphi_{\text{recursion}} := \forall x \forall z (\varphi(x, z) \vee \psi(x, z))$  (see Appendix B), where

$\varphi(x, z) := \neg \exists y F(x, y) \wedge \neg \exists y F(z, y)$  and

$$\psi(x, z) := \bigwedge_{\substack{k'_1, k'_2 \in [D]^2 \\ \ell'_1, \ell'_2 \in [D]^2}} \left( \exists y [E_{k'_1, \ell'_1}(x, y) \wedge E_{k'_2, \ell'_2}(y, z)] \rightarrow \right. \\ \left. \bigwedge_{\substack{i, j, i', j' \in [D], k, \ell \in ([D]^2)^2 \\ \text{ROT}_H(k, i) = ((k'_1, k'_2), i') \\ \text{ROT}_H((\ell'_2, \ell'_1), j) = (\ell, j')}} \exists x' \exists z' [F_k(x, x') \wedge F_\ell(z, z') \wedge E_{(i, j), (j', i')}(x', z')] \right).$$

Let  $a, c \in U(A)$ . Assume first that there is  $b \in U(A)$  with  $(a, b) \in F(A)$ . Hence  $A \not\models \varphi(a, c)$ . Since  $\varphi_{\text{recursion}} := \forall x \forall z (\varphi(x, z) \vee \psi(x, z))$  we aim to prove  $A \models \psi(a, c)$ . By construction of  $\rho_k$ , there is an  $i \in I_k$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d, 2, \sigma}^i$ . Therefore there is a structure  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Let  $f$  be an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Since  $b \in \mathcal{N}_2^A(a)$ , we get that  $f(b)$  is defined. Since  $f$  is an isomorphism mapping  $a$  onto  $\tilde{a}$ , we have that  $(a, b) \in F(A)$  implies that  $(\tilde{a}, f(b)) \in F(\tilde{A})$ . Hence  $\tilde{A} \not\models \varphi(\tilde{a}, \tilde{c})$ , for every  $\tilde{c} \in U(\tilde{A})$ . But since  $\tilde{A} \models \varphi_{\text{recursion}}$ , as  $\tilde{A} \models \varphi_{\mathbb{Z}}$ , this shows that  $\tilde{A} \models \psi(\tilde{a}, \tilde{c})$  for every  $\tilde{c} \in U(\tilde{A})$ .

Let  $k'_1, k'_2 \in [D]^2$  and  $\ell'_1, \ell'_2 \in [D]^2$  be indices such that there is  $b' \in U(A)$  with  $(a, b') \in E_{k'_1, \ell'_1}(A)$  and  $(b', c) \in E_{k'_2, \ell'_2}(A)$ . Since  $b', c \in \mathcal{N}_2^A(a)$ , by assumption we get that  $f(b')$  and  $f(c)$  are defined. Furthermore,  $(a, b') \in E_{k'_1, \ell'_1}(A)$  and  $(b', c) \in E_{k'_2, \ell'_2}(A)$  imply that  $(\tilde{a}, f(b')) \in E_{k'_1, \ell'_1}(\tilde{A})$  and  $(f(b'), f(c)) \in E_{k'_2, \ell'_2}(\tilde{A})$ , since  $f$  is an isomorphism mapping  $a$  onto  $\tilde{a}$ . We proved in the previous paragraph that  $\tilde{A} \models \psi(\tilde{a}, f(c))$ . Hence we can conclude that for all indices  $i, j, i', j' \in [D]$ ,  $k, \ell \in ([D]^2)^2$  for which  $\text{ROT}_H(k, i) = ((k'_1, k'_2), i')$  and  $\text{ROT}_H((\ell'_2, \ell'_1), j) = (\ell, j')$ , there are elements  $\tilde{a}', \tilde{c}' \in U(\tilde{A})$  such that  $(\tilde{a}, \tilde{a}') \in F_k(\tilde{A})$ ,  $(f(c), \tilde{c}') \in F_\ell(\tilde{A})$ , and  $(\tilde{a}', \tilde{c}') \in E_{(i, j), (j', i')}(\tilde{A})$ . Since  $\tilde{a}', \tilde{c}' \in \mathcal{N}_2^{\tilde{A}}(\tilde{a})$ , we get that  $a' := f^{-1}(\tilde{a}')$  and  $c' := f^{-1}(\tilde{c}')$  are defined. Furthermore, we get that  $(a, a') \in F_k(A)$ ,  $(c, c') \in F_\ell(A)$  and  $(a', c') \in E_{(i, j), (j', i')}(A)$ . This proves that  $A \models \psi(a, c)$ .

In the case that there is  $b \in U(A)$  with  $(c, b) \in F(A)$ , we can prove similarly that  $A \models \psi(a, c)$ , by considering that there exist  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{c} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), c) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{c}), \tilde{c})$  by construction of  $\rho_k$ . Finally if there is no  $b \in U(A)$  such that  $(a, b) \in F(A)$  or  $(c, b) \in F(A)$  then  $A \models \varphi(a, c)$ . Since this covers every case we get that  $A \models \varphi_{\text{recursion}}$ .  $\triangleleft$

Assume  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k}$ . As proved in Claims 30, 31, 32 and 21 this implies that  $A \models \varphi'_{\text{tree}}$ ,  $A \models \varphi_{\text{rotationMap}}$ ,  $A \models \varphi_{\text{base}}$  and  $A \models \varphi_{\text{recursion}}$ . Since  $\varphi'_{\mathbb{Z}}$  is a conjunction of these formulas, we get  $A \models \varphi'_{\mathbb{Z}}$  and hence  $A \in \mathcal{P}'_{\mathbb{Z}}$ .  $\blacktriangleleft$

## 4.2 A local reduction from relational structures to graphs

In this section we will define our graph property  $\mathcal{P}_{\text{graph}}$  by giving a reduction from the property  $\mathcal{P}'_{\mathbb{Z}}$  and argue that  $\mathcal{P}_{\text{graph}}$  is GSF-local while not testable. To do so, we show that this reduction is ‘local’ which preserves the testability of these two properties.

### Local reduction

We first introduce the following notion of local reduction between two property testing models. In the following, when the context is clear, we will use  $\mathcal{C}$  to denote both a class of structure



and the corresponding property testing model, which can be either the bounded-degree model for graphs or bounded-degree model for relational structures.

► **Definition 22** (Local reduction). *Let  $\mathcal{C}, \mathcal{C}'$  be two property testing models and let  $\mathcal{P} \subseteq \mathcal{C}$ ,  $\mathcal{P}' \subseteq \mathcal{C}'$  be two properties. We say that a function  $f : \mathcal{C} \rightarrow \mathcal{C}'$  is a local reduction from  $\mathcal{P}$  to  $\mathcal{P}'$  if there are constants  $c_1, c_2 \in \mathbb{N}_{\geq 1}$  such that for every  $X \in \mathcal{C}$  the following properties hold.*

1. *If  $X \in \mathcal{P}$  then  $f(X) \in \mathcal{P}'$ .*
2. *If  $X$  is  $\epsilon$ -far from  $\mathcal{P}$  then  $f(X)$  is  $(\epsilon/c_1)$ -far from  $\mathcal{P}'$ .*
3. *For every query to  $f(X)$  we can adaptively<sup>2</sup> compute  $c_2$  queries such that the answer to the query to  $f(X)$  can be computed from the answers to the  $c_2$  queries to  $X$ .*

The following lemma is known.

► **Lemma 23** (Theorem 7.14 in [9]). *Let  $\mathcal{C}, \mathcal{C}'$  be two property testing models,  $\mathcal{P} \subseteq \mathcal{C}$ ,  $\mathcal{P}' \subseteq \mathcal{C}'$  be two properties and  $f$  a local reduction from  $\mathcal{P}$  to  $\mathcal{P}'$ . If  $\mathcal{P}'$  is testable then so is  $\mathcal{P}$ .*

### Construction of the graph property

Now we construct a property  $\mathcal{P}_{\text{graph}}$  from the property  $\mathcal{P}'_{\mathbb{Z}}$ . We obtain this graph property as  $f(\mathcal{P}'_{\mathbb{Z}})$  by defining a map  $f : \mathcal{C}_{\sigma,d} \rightarrow \mathcal{C}_d$ . To define  $f$  we introduce a distinct arrow-graph gadget for every relation in  $\sigma$  (i.e. for every edge colour). The map  $f$  then replaces every tuple in a certain relation (every coloured edge) by the respective arrow-graph gadget. We further prove that this replacement operation defines a local reduction  $f$  from  $\mathcal{P}'_{\mathbb{Z}}$  to  $\mathcal{P}_{\text{graph}}$ . Recall that a local reduction is a function maintaining distance that can be simulated locally by queries. Since by Lemma 23 local reductions preserve testability, we use the local reduction from  $\mathcal{P}'_{\mathbb{Z}}$  to  $\mathcal{P}_{\text{graph}}$  to obtain non-testability of the property  $\mathcal{P}_{\text{graph}}$  from the non-testability of  $\mathcal{P}'_{\mathbb{Z}}$ . We will now define  $f$  formally.

Let  $\ell$  be the number of relations (the number of edge colours) in  $\sigma$ . We first introduce the different types of arrow-graph gadgets we need to define the local reduction. For  $1 \leq k \leq \ell$ , we let  $H_k$  be the graph with vertex set  $V(H_k) := \{a_1, \dots, a_{2\ell+2}, b_1, b_2\}$  and edge set  $E(H_k) := \{\{a_i, a_{i+1}\} \mid 1 \leq i \leq 2\ell+1\} \cup \{\{a_{\ell+1+k}, b_j\} \mid j \in \{1, 2\}\}$ . We call  $H_k$  a  $k$ -arrow. For any graph  $G$  and vertices  $v, w \in V(G)$ , we say that there is a  $k$ -arrow from  $v$  to  $w$ , denoted  $v \xrightarrow{k} w$ , if there are  $2\ell+2$  vertices  $v_2, \dots, v_{2\ell+1}, w_1, w_2 \in V(G)$  and an isomorphism  $g : H_k \rightarrow \mathcal{N}_1^G(v_2, \dots, v_{2\ell+1}, w_1, w_2)$  such that  $g(a_1) = v$  and  $g(a_{2\ell+2}) = w$ . We now define a second arrow gadget. For  $1 \leq k \leq \ell$ , we let  $L_k$  be the graph with vertex set  $V(L_k) := \{a_1, \dots, a_{\ell+1}, b\}$  and edge set  $E(L_k) := \{\{a_i, a_{i+1}\} \mid 1 \leq i \leq \ell\} \cup \{\{a_k, b\}\}$ . We call  $L_k$  a  $k$ -loop. For any graph  $G$  and vertex  $v \in V(G)$ , we say that there is a  $k$ -loop at  $v$ , denoted  $v \xrightarrow{k} v$ , if there are  $\ell+1$  vertices  $v_1, \dots, v_\ell, w \in V(G)$  and an isomorphism  $g : L_k \rightarrow \mathcal{N}_1^G(v_1, \dots, v_\ell, w)$  such that  $g(a_{\ell+1}) = v$ . Finally we let  $H_\perp$  be the graph with vertex set  $V(H_\perp) := \{a_1, \dots, a_{\ell+1}, b\}$  and edge set  $E(H_\perp) := \{\{a_i, a_{i+1}\} \mid 1 \leq i \leq \ell\} \cup \{\{a_i, b\} \mid i \in \{1, 2\}\}$ . We call  $H_\perp$  a non-arrow. For any graph  $G$  and vertex  $v \in V(G)$ , we say that there is a non-arrow at  $v$ , denoted  $v \not\xrightarrow{\cdot} v$ , if there are  $\ell+1$  vertices  $v_1, \dots, v_\ell, w \in V(G)$  and an isomorphism  $g : H_\perp \rightarrow \mathcal{N}_1^G(v_1, \dots, v_\ell, w)$  such that  $g(a_{\ell+1}) = v$ .

We now define a function  $f : \mathcal{C}_{\sigma,d} \rightarrow \mathcal{C}_d$  by  $f(A) := G_A$ , where  $G_A$  is the graph on vertex

<sup>2</sup> By adaptively computing queries we mean that the selection of the next query may depend on the answer to the previous query.

34:16 GSF-locality is not sufficient for proximity-oblivious testing

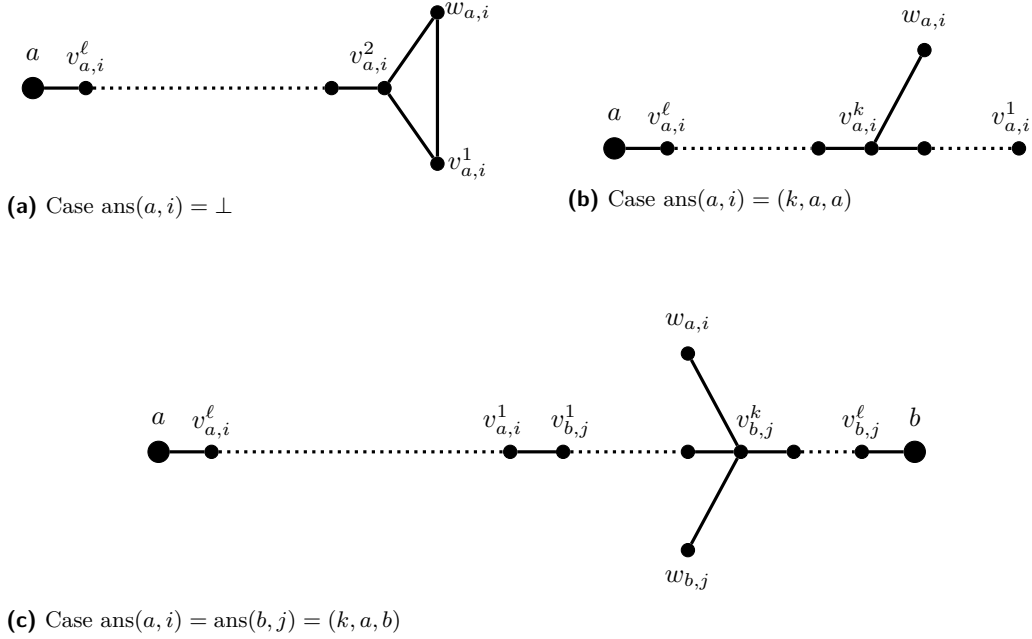


Figure 3 Different types of arrows in  $G_A$ .

set  $V(G_A) := U(A) \cup \{v_{a,i}^k, w_{a,i} \mid 1 \leq i \leq d, a \in U(A), 1 \leq k \leq \ell\}$  and edge set

$$\begin{aligned}
 E(G_A) := & \left\{ \{a, v_{a,i}^\ell\} \mid a \in U(A), 1 \leq i \leq d \right\} \\
 & \cup \left\{ \{v_{a,i}^k, v_{a,i}^{k+1}\} \mid 1 \leq k \leq \ell - 1, a \in U(A), 1 \leq i \leq d \right\} \\
 & \cup \left\{ \{v_{b,j}^k, w_{b,j}\}, \{v_{b,j}^k, w_{a,i}\}, \{v_{a,i}^\ell, v_{b,j}^\ell\} \mid a \neq b, \text{ans}(a, i) = \text{ans}(b, j) = (k, a, b) \right\} \\
 & \cup \left\{ \{v_{a,i}^k, w_{a,i}\} \mid \text{ans}(a, i) = (k, a, a) \right\} \cup \left\{ \{v_{a,i}^1, w_{a,i}\}, \{v_{a,i}^2, w_{a,i}\} \mid \text{ans}(a, i) = \perp \right\},
 \end{aligned}$$

where  $\text{ans}(a, i) = (k, a, b)$  denotes that the  $i$ -th tuple of  $a$  is  $(a, b)$  and is in the  $k$ -th relation. Hence  $G_A$  is defined in such a way that if  $(a, b)$  is a tuple in the  $k$ -th relation of  $\sigma$  in  $A$ , then  $a \xrightarrow{k} b$  in  $G_A$ , and  $a$  has a non-arrow for every  $i$  satisfying that  $\text{ans}(a, i) = \perp$  for every  $k$ . For illustration see Figure 3.

Now we define property  $\mathcal{P}_{\text{graph}} := \{f(A) \mid A \in \mathcal{P}'_{(\mathbb{Z})}\} \subseteq \mathcal{C}_d$ .

► **Lemma 24.** *The map  $f$  is a local reduction from  $\mathcal{P}'_{(\mathbb{Z})}$  to  $\mathcal{P}_{\text{graph}}$ .*

**Proof.** First note that for any  $A \in \mathcal{P}'_{(\mathbb{Z})}$ , we have that  $f(A) \in \mathcal{P}_{\text{graph}}$  by definition.

Now let  $c_1 = 2d + 2d^2\ell$ . We prove that if  $A \in \mathcal{C}_{\sigma,d}$  is  $\epsilon$ -far from  $\mathcal{P}'_{(\mathbb{Z})}$  then  $f(A)$  is  $\epsilon/c_1$ -far from  $\mathcal{P}_{\text{graph}}$  by contraposition. Therefore assume that  $f(A) =: G_A$  is not  $\epsilon/c_1$ -far from  $\mathcal{P}_{\text{graph}}$  for some  $A \in \mathcal{C}_{\sigma,d}$ . Then there is a set  $E \subseteq \{e \subseteq V(G_A) \mid |e| = 2\}$  of size at most  $\epsilon d |V(G_A)| / c_1$ , and a graph  $G \in \mathcal{P}_{\text{graph}}$  such that  $G$  is obtained from  $G_A$  by modifying the tuples in  $E$ . By definition of  $\mathcal{P}_{\text{graph}}$ , there is a structure  $A_G \in \mathcal{P}'_{(\mathbb{Z})}$  such that  $f(A_G) = G$ . First note that  $|U(A_G)| = |U(A)|$ , as  $(1 + d\ell)|U(A)| = |V(G_A)| = |V(G)| = (1 + d\ell)|U(A_G)|$ . Hence there must be a set  $R$  of tuples that need to be modified to make  $A$  isomorphic to  $A_G$ . First note that  $R$  cannot contain a tuple  $(a, b)$  where  $\{a, v_{a,i}^k, w_{a,i}, b, v_{b,i}^k, w_{b,i} \mid 1 \leq i \leq d, 1 \leq k \leq \ell\} \cap e = \emptyset$  for every  $e \in E$ . This is because if  $(a, b)$  is a tuple in  $A$ , then  $a \xrightarrow{k} b$  for some  $k$  in

$G_A$ . But since  $\{a, v_{a,i}^k, w_{a,i}, b, v_{b,i}^k, w_{b,i} \mid 1 \leq i \leq d, 1 \leq k \leq \ell\} \cap e = \emptyset$  for every  $e \in E$ , we have that  $a \xrightarrow{k} b$  in  $G$ . But then  $(a, b)$  must be a tuple in  $A_G$ , and hence  $(a, b)$  cannot be in  $R$ . The same argument works when assuming that  $(a, b)$  is a tuple in  $A_G$ . Since for every  $e \in E$ , there are at most  $2d$  tuples  $(a, b)$  such that  $\{a, v_{a,i}^k, w_{a,i}, b, v_{b,i}^k, w_{b,i} \mid 1 \leq i \leq d, 1 \leq k \leq \ell\} \cap e \neq \emptyset$ , we get that

$$|R| \leq 2d\epsilon d|V(G_A)|/c_1 = 2d(1 + d\ell)\epsilon d|U(A)|/c_1 = \epsilon d|U(A)|.$$

Hence  $A$  is not  $\epsilon$ -far to being in  $\mathcal{P}'_{\mathbb{Z}}$ .

Let  $c_2 := d + 1$ . Let  $A \in \mathcal{C}_{\sigma,d}$  and  $G_A := f(A)$ . Note that any  $a \in U(A)$  is adjacent in  $G_A$  to  $v_{a,i}^\ell$ , for every  $1 \leq i \leq d$ . Hence any neighbour query in  $G_A$  to  $a$  can be answered without querying  $A$ . Assume  $v \in \{v_{a,i}^k, w_{a,i} \mid 1 \leq k \leq \ell\}$  for some  $a \in U(A)$  and some  $1 \leq i \leq d$ . Then we can determine all neighbours of  $v$  by querying  $(a, i)$  and further if  $\text{ans}(a, i) \neq \perp$  and  $\text{ans}(a, i) = (k, a, b)$ , then we need to query  $(b, j)$  for every  $1 \leq j \leq d$ . Hence we can determine the answer to any query to  $G_A$  by making  $c_2$  queries to  $A$ . This proves that  $f$  is a local reduction from  $\mathcal{P}'_{\mathbb{Z}}$  to  $\mathcal{P}_{\text{graph}}$ .  $\blacktriangleleft$

We remark that  $\mathcal{P}_{\text{graph}}$  is a simpler version of the simple graph property defined in [1] where extra care had to be taken to define degree-regular graphs.

### 4.3 The graph property is GSF-local

Let  $\mathcal{P}_{\text{graph}}$  be the graph property as defined in Section 4.2. We now show that  $\mathcal{P}_{\text{graph}}$  is GSF-local.

► **Lemma 25.** *The graph property  $\mathcal{P}_{\text{graph}}$  is GSF-local.*

**Proof.** For this we will prove that  $\mathcal{P}_{\text{graph}}$  is equal to a finite union of properties defined by 0-profiles, and then use Theorem 11 to prove that  $\mathcal{P}_{\text{graph}}$  is GSF-local. We define the 0-profiles for  $\mathcal{P}_{\text{graph}}$  in a very similar way to the relational structure case, and then use the description of  $\mathcal{P}'_{\mathbb{Z}}$  by 0-profiles shown in Lemma 20. To this end, assume that the  $4\ell + 2$ -types  $\tau_{d,4\ell+2}^1, \dots, \tau_{d,4\ell+2}^{n_{d,4\ell+2}}$  are ordered in such a way that  $(\mathcal{N}_{4\ell+2}^{f(B)}(b), b) \in \tau_{d,4\ell+2}^k$ , for every  $k \in \{1, \dots, m\}$  and  $(B, b) \in \tau_{d,2,\sigma}^k$ , where  $m$  is the number of parts of the partition of  $\mathcal{P}_{\mathbb{Z}}$  defined in Subsection 4.1. For  $k \in \{1, \dots, m\}$ , let  $\hat{I}_k$  be the set of indices  $i$  such that there is  $A \in \mathcal{P}_k$ , and  $v \in V(f(A))$  for which  $(\mathcal{N}_{4\ell+2}^{f(A)}(v), v) \in \tau_{d,4\ell+2}^i$ . Let  $\hat{\rho}_k : \{1, \dots, n_{d,4\ell+2}\} \rightarrow \mathfrak{I}_0$  be defined by

$$\hat{\rho}_k(i) := \begin{cases} [0, 1] & \text{if } i = k, \\ [0, \infty) & \text{if } i \in \hat{I}_k \setminus \{k\}, \\ [0, 0] & \text{otherwise.} \end{cases}$$

► **Claim 26.** It holds that  $\mathcal{P}_{\text{graph}} = \bigcup_{1 \leq k \leq m} \mathcal{P}_{\hat{\rho}_k}$ .

**Proof.** First we prove  $\mathcal{P}_{\text{graph}} \subseteq \bigcup_{1 \leq k \leq m} \mathcal{P}_{\hat{\rho}_k}$ . Assume  $G \in \mathcal{P}_{\text{graph}}$  and let  $A \in \mathcal{P}'_{\mathbb{Z}}$  be a structure such that  $G = f(A)$ . If  $A = A_\emptyset$  then clearly  $G \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\hat{\rho}_k}$ . Hence assume  $A \neq A_\emptyset$ . Then  $A \in \mathcal{P}_k$  for some  $k \in \{1, \dots, m\}$ . By the construction of  $\hat{I}_k$  we know that for every  $v \in V(G)$  we have  $(\mathcal{N}_{4\ell+2}^G(v), v) \in \tau_{d,4\ell+2}^i$  for some  $i \in \hat{I}_k$ . Furthermore, since  $A \in \mathcal{P}_k$  there is at most one  $a \in U(A)$  with  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^k$ . This implies directly

that there can be at most one vertex  $v \in V(G)$  with  $(\mathcal{N}_{4\ell+2}^G(v), v) \in \tau_{d,4\ell+2}^k$  and hence  $G \in \mathcal{P}_{\hat{\rho}}$ .

Now we prove that  $\bigcup_{1 \leq k \leq m} \mathcal{P}_{\hat{\rho}_k} \subseteq \mathcal{P}_{\text{graph}}$ . Let  $G \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\hat{\rho}_k}$  and let  $k \in \{1, \dots, m\}$  be an index such that  $G \in \mathcal{P}_{\hat{\rho}_k}$ .

First note that every model of  $\varphi_{\mathbb{Z}}$  is  $d$  regular for some large  $d$ . Then for any  $A \models \varphi_{\mathbb{Z}}$ , every vertex in  $f(A)$  has either degree  $\leq 4$  or degree  $d$ . Since every structure in  $\mathcal{P}'_{\mathbb{Z}}$  apart from the empty structure  $A_\emptyset$  is a model of  $\varphi_{\mathbb{Z}}$ , this implies that every vertex in any graph  $G' \in \mathcal{P}_{\text{graph}}$  has degree  $\leq 4$  or degree  $d$ . Since for every  $i$  for which  $\hat{\rho}(i) \neq [0, 0]$ , there is a graph  $G' \in \mathcal{P}_{\text{graph}}$  and  $v \in V(G')$  such that  $(\mathcal{N}_{4\ell+2}^{G'}(v), v) \in \tau_{d,4\ell+2}^i$ , we get that every vertex in  $G$  has to have degree  $\leq 4$  or degree  $d$ . Using this argument further, we get that every vertex  $v \in V(G)$  of degree  $\leq 4$  has to be contained in the  $(\ell + 1)$ -neighbourhood of a vertex of degree  $d$ , and that the  $(2\ell + 1)$ -neighbourhood of every vertex  $v \in V(G)$  of degree  $d$  is the union of  $k$ -arrows,  $k$ -loops and non-arrows which are disjoint apart from their endpoints. Hence there is a  $\sigma$ -structure  $A$  such that  $f(A) \cong G$ . Let  $g$  be an isomorphism from  $f(A)$  to  $G$ .

Now we argue that  $A \in \mathcal{P}_{\rho_k}$ . First assume that there are two elements  $a, b$  with  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^k$  and  $(\mathcal{N}_2^A(b), b) \in \tau_{d,2,\sigma}^k$ . By definition, we get that  $(\mathcal{N}_{4\ell+2}^{f(A)}(a), a) \in \tau_{d,4\ell+2}^k$  and  $(\mathcal{N}_{4\ell+2}^{f(A)}(b), b) \in \tau_{d,4\ell+2}^k$ . Since  $g$  is an isomorphism, the restriction of  $g$  to  $\mathcal{N}_{4\ell+2}^{f(A)}(a)$  must be an isomorphism from  $\mathcal{N}_{4\ell+2}^{f(A)}(a)$  to  $\mathcal{N}_{4\ell+2}^G(g(a))$ , and hence  $(\mathcal{N}_{4\ell+2}^G(g(a)), g(a)) \cong (\mathcal{N}_{4\ell+2}^{f(A)}(a), a) \in \tau_{d,4\ell+2}^k$ . But the same holds for the  $(4\ell + 2)$ -ball of  $g(b)$ , and hence we contradict the assumption that  $G \in \mathcal{P}_{\hat{\rho}_k}$  since  $\hat{\rho}_k(k) = [0, 1]$ . Let us further assume that there is an  $a \in U(A)$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^i$  for some  $i \notin I_k$ . Let  $j$  be the index such that  $(\mathcal{N}_{4\ell+2}^{f(A)}(a), a) \in \tau_{d,4\ell+2}^j$ . Additionally note that  $a$  must have degree  $d$  in  $f(A)$  by construction of  $f$ . As  $g$  is an isomorphism, we get that  $(\mathcal{N}_{4\ell+2}^G(g(a)), g(a)) \in \tau_{d,4\ell+2}^j$ , and  $g(a)$  has degree  $d$ . But then by construction of  $\hat{\rho}_k$ , there must be  $G' \in \mathcal{P}_{\text{graph}}$ , and a vertex  $v \in V(G')$  of degree  $d$  such that  $(\mathcal{N}_{4\ell+2}^{G'}(v), v) \in \tau_{d,4\ell+2}^j$ . By construction of  $\mathcal{P}_{\text{graph}}$ , there is a structure  $A' \in \mathcal{P}'_{\mathbb{Z}}$  such that  $f(A') = G'$ . Since  $v$  has degree  $d$ , it must be an element in  $A'$ . Furthermore  $(\mathcal{N}_2^{A'}(v), v) \in \tau_{d,2,\sigma}^i$  by choice of  $i$  and  $j$ . Hence  $A' \notin \mathcal{P}_{\rho_k}$ . But this contradicts Lemma 20.

Hence we have shown that  $A \in \mathcal{P}_{\rho_k}$ . Then by Lemma 20  $A \in \mathcal{P}'_{\mathbb{Z}}$ , and by construction  $G \in \mathcal{P}_{\text{graph}}$ .  $\triangleleft$

Since by Claim 26 we can express  $\mathcal{P}_{\text{graph}}$  as a finite union of properties each defined by a 0-profile, Theorem 11 implies that  $\mathcal{P}_{\text{graph}}$  is GSF-local.  $\blacktriangleleft$

#### 4.4 Putting everything together

Now we prove our main theorem.

**Proof of Theorem 1.** Let the property  $\mathcal{P}'_{\mathbb{Z}}$  of relational structures be as defined above. Note that  $\mathcal{P}'_{\mathbb{Z}}$  is not testable, as  $\mathcal{P}_{\mathbb{Z}}$  is not testable [1, Theorem 4.4] and  $\mathcal{P}'_{\mathbb{Z}}$  only differs from  $\mathcal{P}_{\mathbb{Z}}$  by the empty structure. By Lemma 24 and Lemma 23, the graph property  $\mathcal{P}_{\text{graph}}$  that is locally reduced from  $\mathcal{P}'_{\mathbb{Z}}$  is not testable. Lemma 25 shows that  $\mathcal{P}_{\text{graph}}$  is also a GSF-local property. Hence there exists a GSF-local property of bounded-degree graphs which is not testable. Furthermore, since having a POT implies being testable, this proves that there is a GSF-local property which has no POT. By Theorem 6 this implies that not all GSF-local properties are non-propagating.  $\blacktriangleleft$

---

References

---

- 1 Isolde Adler, Noleen Köhler, and Pan Peng. On testability of first-order properties in bounded-degree graphs. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1578–1597. SIAM, 2021.
- 2 Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: it’s all about regularity. *SIAM Journal on Computing*, 39(1):143–167, 2009.
- 3 Itai Benjamini, Oded Schramm, and Asaf Shapira. Every minor-closed property of sparse graphs is testable. *Advances in mathematics*, 223(6):2200–2218, 2010.
- 4 Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of computer and system sciences*, 47(3):549–595, 1993.
- 5 Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite model theory*. Perspectives in Mathematical Logic. Springer, 1995.
- 6 Hendrik Fichtenberger, Pan Peng, and Christian Sohler. Every testable (infinite) property of bounded-degree graphs contains an infinite hyperfinite subproperty. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 714–726. Society for Industrial and Applied Mathematics, 2019.
- 7 Sebastian Forster, Danupon Nanongkai, Thatchaphol Saranurak, Liu Yang, and Sorrachai Yingchareonthawornchai. Computing and testing small connectivity in near-linear time and queries via fast local cut algorithms. *SODA*, 2020.
- 8 Haim Gaifman. On local and non-local properties. In *Studies in Logic and the Foundations of Mathematics*, volume 107, pages 105–135. Elsevier, 1982.
- 9 Oded Goldreich. *Introduction to property testing*. Cambridge University Press, 2017.
- 10 Oded Goldreich, Shari Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM (JACM)*, 45(4):653–750, 1998.
- 11 Oded Goldreich and Tali Kaufman. Proximity oblivious testing and the role of invariances. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 173–190. Springer, 2011.
- 12 Oded Goldreich and Dana Ron. Property testing in bounded degree graphs. *Algorithmica*, 32(2):302–343, 2002. URL: <http://dx.doi.org/10.1007/s00453-001-0078-7>, doi:10.1007/s00453-001-0078-7.
- 13 Oded Goldreich and Dana Ron. On proximity-oblivious testing. *SIAM Journal on Computing*, 40(2):534–566, 2011. Preliminary version appeared at *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC 2009)*.
- 14 Oded Goldreich and Igor Shinkar. Two-sided error proximity oblivious testing. *Random Structures & Algorithms*, 48(2):341–383, 2016.
- 15 William Hanf. *The Theory of Models*, chapter Model-theoretic methods in the study of elementary logic, pages 132–145. North Holland, 1965.
- 16 Avinatan Hassidim, Jonathan A Kelner, Huy N Nguyen, and Krzysztof Onak. Local graph partitions for approximation and testing. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 22–31. IEEE, 2009.
- 17 Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *BULL. AMER. MATH. SOC.*, 43(4):439–561, 2006.
- 18 Hiro Ito, Areej Khoury, and Ilan Newman. On the characterization of 1-sided error strongly testable graph properties for bounded-degree graphs. *Computational Complexity*, 29(1):1–45, 2020.
- 19 Ken-ichi Kawarabayashi and Yuichi Yoshida. Testing subdivision-freeness: property testing meets structural graph theory. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 437–446. ACM, 2013.
- 20 Akash Kumar, C Seshadhri, and Andrew Stolman. Random walks and forbidden minors ii: a poly( $d \varepsilon^{-1}$ )-query tester for minor-closed properties of bounded degree graphs. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 559–567, 2019.

- 21 László Lovász. *Large Networks and Graph Limits*, volume 60 of *Colloquium Publications*. American Mathematical Society, 2012.
- 22 Ilan Newman and Christian Sohler. Every property of hyperfinite graphs is testable. *SIAM Journal on Computing*, 42(3):1095–1112, 2013.
- 23 Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of mathematics*, pages 157–187, 2002.
- 24 Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25(2):252–271, 1996.
- 25 Yuichi Yoshida and Hiro Ito. Property testing on  $k$ -vertex-connectivity of graphs. *Algorithmica*, 62(3-4):701–712, 2012.

## A

 Formal definitions of property testers and POTs

Now we give the formal definitions of standard property testing and proximity-oblivious testing.

► **Definition 27** ((Standard) property testing). *Let  $\mathcal{P} = \cup_{n \in \mathbb{N}} \mathcal{P}_n$  be a property. An  $\epsilon$ -tester for  $\mathcal{P}_n$  is a probabilistic algorithm which, given query access to a structure  $A \in \mathcal{C}$  with  $n$  vertices/elements,*

- *accepts  $A$  with probability  $2/3$  if  $A \in \mathcal{P}_n$ .*
- *rejects  $A$  with probability  $2/3$  if  $A$  is  $\epsilon$ -far from  $\mathcal{P}_n$ .*

*We say that a property  $\mathcal{P}$  is testable if for every  $n \in \mathbb{N}$  and  $\epsilon \in (0, 1)$ , there exists an  $\epsilon$ -tester for  $\mathcal{P}_n$  that makes at most  $q = q(\epsilon, d)$  queries. We say the property  $\mathcal{P}$  is testable with one-sided error if the  $\epsilon$ -tester always accepts  $A$  if  $A \in \mathcal{P}$ .*

We introduce below the formal definition of proximity-oblivious testers.

► **Definition 28** ((One-sided error) proximity-oblivious testing). *Let  $\mathcal{P} = \cup_{n \in \mathbb{N}} \mathcal{P}_n$  be a property. Let  $\eta : (0, 1] \rightarrow (0, 1]$  be a monotone function. A proximity-oblivious tester (POT) with detection probability  $\eta$  for  $\mathcal{P}_n$  is a probabilistic algorithm which, given query access to a structure  $A \in \mathcal{C}$  with  $n$  vertices/elements,*

- *accepts  $A$  with probability 1 if  $A \in \mathcal{P}_n$ .*
- *rejects  $A$  with probability at least  $\eta(\text{dist}(A, \mathcal{P}_n))$  if  $A \notin \mathcal{P}_n$ , where  $\text{dist}(A, \mathcal{P}_n)$  is the minimum fraction of different edges between  $A$  and any other  $A' \in \mathcal{P}_n$ .*

*We say that a property  $\mathcal{P}$  is proximity-oblivious testable if for every  $n \in \mathbb{N}$ , there exists a POT for  $\mathcal{P}_n$  of constant query complexity with detection probability  $\eta$ .*

## B

 The FO formula

For the construction of the formula  $\varphi_{\mathbb{Z}}$  we use a recursively defined sequence  $(G_m)_{m \in \mathbb{N}_{>0}}$  of edge expanders [17, Proposition 9.2]. Using this sequence we define the formula  $\varphi_{\mathbb{Z}}$  in such a way that any model restricted to relation  $F$  forms a rooted complete  $D^4$ -ary tree. Furthermore, the formula enforces that restricted to the vertices of level  $i$  of the tree the relation  $E$  encodes the rotation map of the expander  $G_i$ . The formula  $\varphi_{\mathbb{Z}}$  is the conjunction of the following formulas. For a more detailed explanation and a proof of the precise form of the models of  $\varphi_{\mathbb{Z}}$  see [1].

We use the following formula

$$\varphi_{\text{root}}(x) := \forall y \neg F(y, x),$$

which expresses that vertex  $x$  is a root vertex, i. e. has no incoming  $F$ -edges. We then define the formula  $\varphi_{\text{tree}}$  which expresses that the structure restricted to the relation  $F$  locally looks like a tree. More precisely, the formula expresses that there is precisely one root vertex, that every other vertex has one incoming  $F$ -edge and every vertex either has no  $F$ -children or has precisely  $D^4$   $F$ -children. We furthermore attach a  $R$ -self-loop to the root and  $D^4$   $L$ -self-loops to the leaves. This was important in [1] to make structures degree regular, but is of no relevance to this proof.

$$\begin{aligned} \varphi_{\text{tree}} := & \exists^=1 x \varphi_{\text{root}}(x) \wedge \\ & \forall x \left( (\varphi_{\text{root}}(x) \wedge R(x, x)) \vee (\exists^=1 y F(y, x) \wedge \neg \exists y R(x, y) \wedge \neg \exists y R(y, x)) \right) \wedge \\ & \forall x \left( \left[ \neg \exists y F(x, y) \wedge \bigwedge_{k \in ([D]^2)^2} L_k(x, x) \wedge \forall y (y \neq x \rightarrow \bigwedge_{k \in ([D]^2)^2} \neg L_k(x, y) \wedge \right. \right. \\ & \quad \left. \bigwedge_{k \in ([D]^2)^2} \neg L_k(y, x) \right] \vee \left[ \neg \exists y \bigvee_{k \in ([D]^2)^2} (L_k(x, y) \vee L_k(y, x)) \wedge \right. \\ & \quad \left. \bigwedge_{k \in ([D]^2)^2} \exists y_k \left( x \neq y_k \wedge F_k(x, y_k) \wedge \left( \bigwedge_{\substack{k' \in ([D]^2)^2 \\ k' \neq k}} \neg F_{k'}(x, y_k) \right) \wedge \forall y (y \neq y_k \rightarrow \neg F_k(x, y)) \right) \right] \right). \end{aligned}$$

We define formula  $\varphi_{\text{rotationMap}}$  which expresses that the edge relations restricted to the relations  $E$  encode a rotation map.

$$\begin{aligned} \varphi_{\text{rotationMap}} := & \forall x \forall y \left( \bigwedge_{i, j \in [D]^2} (E_{i, j}(x, y) \rightarrow E_{j, i}(y, x)) \right) \wedge \\ & \forall x \left( \bigwedge_{i \in [D]^2} \left( \bigvee_{j \in [D]^2} (\exists^=1 y E_{i, j}(x, y) \wedge \bigwedge_{\substack{j' \in [D]^2 \\ j' \neq j}} \neg \exists y E_{i, j'}(x, y)) \right) \right) \end{aligned}$$

The formula  $\varphi_{\text{base}}$  expresses that the children of the root vertex form the basis of the recursive construction of expanders. The basis of the recursive construction is the square of some  $D$  regular graph  $H$  on  $D^4$  vertices with edge expansion ratio  $1/4$ . Explicit constructions of graphs with such properties are given in [23]. We assume that this graph is given by a rotation map  $\text{ROT}_H$ , which is an encoding of  $H$ .

$$\begin{aligned} \varphi_{\text{base}} := & \forall x \left( \varphi_{\text{root}}(x) \rightarrow \left[ \bigwedge_{i, j \in [D]^2} \left( E_{i, j}(x, x) \wedge \forall y (x \neq y \rightarrow (\neg E_{i, j}(x, y) \wedge \neg E_{i, j}(y, x))) \right) \right] \wedge \right. \\ & \left. \bigwedge_{\substack{\text{ROT}_{H^2}(k, i) = (k', i') \\ k, k' \in ([D]^2)^2 \\ i, i' \in [D]^2}} \exists y \exists y' (F_k(x, y) \wedge F_{k'}(x, y') \wedge E_{i, i'}(y, y')) \right] \end{aligned}$$

We define the formula  $\varphi_{\text{recursion}}$  which expresses the recursive construction of the sequence



$(G_m)_{m \in \mathbb{N}_{>0}}$ . This formula also depends on the base graph  $H$ .

$$\begin{aligned} \varphi_{\text{recursion}} := & \forall x \forall z \left[ \left( \neg \exists y F(x, y) \wedge \neg \exists y F(z, y) \right) \vee \right. \\ & \bigwedge_{\substack{k'_1, k'_2 \in [D]^2 \\ \ell'_1, \ell'_2 \in [D]^2}} \left( \exists y [E_{k'_1, \ell'_1}(x, y) \wedge E_{k'_2, \ell'_2}(y, z)] \rightarrow \right. \\ & \left. \left. \bigwedge_{\substack{i, j, i', j' \in [D], k, \ell \in ([D]^2)^2 \\ \text{ROT}_H(k, i) = ((k'_1, k'_2), i') \\ \text{ROT}_H((\ell'_2, \ell'_1), j) = (\ell, j')}} \exists x' \exists z' [F_k(x, x') \wedge F_\ell(z, z') \wedge E_{(i, j), (j', i')}(x', z')] \right] \right] \end{aligned}$$

## C Deferred proofs from Section 3

**Proof of Lemma 9.** For the first direction assume  $\varphi$  is an FO-sentence. Then by Hanf's Theorem (Theorem 2) there is a sentence  $\psi$  in Hanf normal form such that  $\mathcal{P}_\varphi = \mathcal{P}_\psi$ .

We will first convert  $\psi$  into a sentence in Hanf normal form where every Hanf sentence appearing has the same locality radius. Let  $r \in \mathbb{N}$  be the maximum locality radius appearing in  $\psi$ , and let  $\varphi_\tau^{\geq m} := \exists^{\geq m} x \varphi_\tau(x)$  be a Hanf sentence, where  $\tau$  is an  $r'$ -type for some  $r' \leq r$ . Let  $\tau_1, \dots, \tau_k$  be a list of all  $r$ -types of bounded degree  $d$  for which  $(\mathcal{N}_{r'}^B(b), b) \in \tau$  for  $(B, b) \in \tau_i$ , for every  $1 \leq i \leq k$ . Let  $\Pi$  be the set of all partitions of  $m$  into  $k$  parts. Let

$$\tilde{\varphi}_\tau^{\geq m} := \bigvee_{(m_1, \dots, m_k) \in \Pi} \bigwedge_{i=1}^k \exists^{\geq m_i} x \varphi_{\tau_i}(x).$$

▷ **Claim 29.**  $\varphi_\tau^{\geq m}$  is  $d$ -equivalent to  $\tilde{\varphi}_\tau^{\geq m}$ .

**Proof.** Assume that  $A \in \mathcal{C}_d$  satisfies  $\varphi_\tau^{\geq m}$ , and assume that  $a_1, \dots, a_m$  are  $m$  distinct elements with  $(\mathcal{N}_{r'}^A(a_j), a_j) \in \tau$ , for every  $1 \leq j \leq m$ . Let  $\tilde{\tau}_j$  be the  $r$ -type for which  $(\mathcal{N}_r^A(a_j), a_j) \in \tilde{\tau}_j$ . By choice of  $\tau_1, \dots, \tau_k$ , we get that there are indices  $i_1, \dots, i_m$  such that  $\tilde{\tau}_j = \tau_{i_j}$ . For  $i \in \{1, \dots, k\}$  let  $m_i = |\{j \in \{1, \dots, m\} \mid i_j = i\}|$ . Hence  $A \models \bigwedge_{i=1}^k \exists^{\geq m_i} x \varphi_{\tau_i}(x)$  and since additionally  $(m_1, \dots, m_k) \in \Pi$  this implies  $A \models \tilde{\varphi}_\tau^{\geq m}$ .

On the other hand, let  $A \in \mathcal{C}_d$  satisfy  $\tilde{\varphi}_\tau^{\geq m}$ , and let  $(m_1, \dots, m_k) \in \Pi$  be a partition of  $m$  such that  $A \models \bigwedge_{i=1}^k \exists^{\geq m_i} x \varphi_{\tau_i}(x)$ . For every  $1 \leq i \leq k$ , let  $a_1^i, \dots, a_{m_i}^i$  be  $m_i$  distinct elements such that  $(\mathcal{N}_r^A(a_j^i), a_j^i) \in \tau_i$ , for every  $1 \leq j \leq m_i$ . By choice of  $\tau_1, \dots, \tau_k$ , we get that  $(\mathcal{N}_{r'}^A(a_j^i), a_j^i) \in \tau$ , for every pair  $1 \leq i \leq k$ ,  $1 \leq j \leq m_i$ . But since  $m_1 + \dots + m_k = m$  this implies that  $A \models \varphi_\tau^{\geq m}$ . This proves that  $\varphi_\tau^{\geq m}$  and  $\tilde{\varphi}_\tau^{\geq m}$  are  $d$ -equivalent. ◁

Let  $\psi'$  be the formula in which every Hanf-sentence  $\varphi_\tau^{\geq m}$  for which  $\tau$  is an  $r'$ -type for some  $r' < r$  gets replaced by  $\tilde{\varphi}_\tau^{\geq m}$ . By a simple inductive argument using Claim 29, we get that  $\psi$  is  $d$ -equivalent to  $\psi'$ , and hence  $\mathcal{P}_\varphi = \mathcal{P}_\psi = \mathcal{P}_{\psi'}$ . Furthermore since  $\tilde{\varphi}_\tau^{\geq m}$  is a Boolean combination of Hanf-sentences for every  $\varphi_\tau^{\geq m}$ , and any Boolean combination of Boolean combinations is a Boolean combination itself,  $\psi'$  is in Hanf normal form. Furthermore, every Hanf-sentence appearing in  $\psi'$  has locality radius  $r$  by construction.

Since any Boolean combination can be converted into disjunctive normal form, we can assume that  $\psi'$  is a disjunction of sentences  $\xi$  of the form

$$\xi = \bigwedge_{j=1}^k \exists^{\geq m_j} x \varphi_{\tau_j}(x) \wedge \bigwedge_{j=k+1}^\ell \neg \exists^{\geq m_j} x \varphi_{\tau_j}(x),$$

where  $\ell \in \mathbb{N}_{\geq 1}$ ,  $1 \leq k \leq \ell$ ,  $m_i \in \mathbb{N}_{\geq 1}$  and  $\tau_i$  is an  $r$ -type for every  $1 \leq i \leq \ell$ . We can further assume that every sentence in the disjunction  $\psi'$  is satisfiable by some  $A \in \mathcal{C}_d$ , as any sentence with no bounded degree  $d$  model can be removed from  $\psi'$ .

Let  $\tilde{\tau}_1, \dots, \tilde{\tau}_t$  be a list of all  $r$ -types of bounded degree  $d$  in the order we fixed. Let  $k_i := \max(\{m_j \mid 1 \leq j \leq k, \tau_j = \tilde{\tau}_i\} \cup \{0\})$  and  $\ell_i := \min(\{m_j \mid k+1 \leq j \leq \ell, \tau_j = \tilde{\tau}_i\} \cup \{\infty\})$  for every  $i \in \{1, \dots, t\}$ . Since  $\xi$  has at least one bounded degree model  $k_i \leq \ell_i$  for every  $i \in \{1, \dots, t\}$ . Let  $\rho : \{1, \dots, t\} \rightarrow \mathcal{T}$  be the neighbourhood profile defined by  $\rho(i) := [k_i, \ell_i]$  if  $\ell_i < \infty$  and  $\rho(i) := [k_i, \ell_i]$  otherwise. Then by construction, we get that  $\mathcal{P}_\rho = \mathcal{P}_\xi$ . Since  $\psi'$  is a disjunction of formulas, each of which defines a property which can be defined by some neighbourhood profile, we get that  $\mathcal{P}_{\psi'}$  must be a finite union of properties defined by some neighbourhood profile.

On the other hand, for every  $r$ -neighbourhood profile  $\rho$  of degree  $d$ ,  $\tau_1, \dots, \tau_t$  a list of all  $r$ -types of bounded degree  $d$  in the order fixed and the formula

$$\varphi_\rho := \bigwedge_{\substack{i \in \{1, \dots, t\}, \\ \rho(i) = [k_i, \ell_i]}} \left( \exists^{\geq k_i} x \varphi_{\tau_i}(x) \wedge \neg \exists^{\geq \ell_i + 1} x \varphi_{\tau_i}(x) \right) \wedge \bigwedge_{\substack{i \in \{1, \dots, t\}, \\ \rho(i) = [k_i, \infty)}} \exists^{\geq k_i} x \varphi_{\tau_i}(x)$$

it clearly holds that  $\mathcal{P}_\rho = \mathcal{P}_{\varphi_\rho}$ . Hence every finite union of properties defined by neighbourhood profiles can be defined by the disjunction of the formulas  $\varphi_\rho$  of all  $\rho$  in the finite union.  $\blacktriangleleft$

## D Deferred proofs from Section 4

Proof of Claim 19. Let  $A, A' \in \mathcal{C}_{\sigma, d}$  such that  $A \models \tilde{\varphi}_{\mathbb{Z}}$  and  $A' \models \tilde{\varphi}_{\mathbb{Z}}$ , where  $\tilde{\varphi}_{\mathbb{Z}}$  was the formula obtained from  $\varphi_{\mathbb{Z}}$  by removing the subformula  $\exists^1 x \varphi_{\text{root}}(x)$ . Our aim is to prove that  $A \cup A' \models \tilde{\varphi}_{\mathbb{Z}}$  where  $A \cup A'$  denotes the disjoint union of  $A$  and  $A'$ . For this we essentially prove that for any two elements  $a \in U(A)$  and  $b \in U(A')$  the formula  $\tilde{\varphi}_{\mathbb{Z}}$  does not require a tuple containing  $a$  and  $b$ .

Let us define formulas

$$\begin{aligned} \varphi &:= \forall x \left( (\varphi_{\text{root}}(x) \wedge R(x, x)) \vee (\exists^1 y F(y, x) \wedge \neg \exists y R(x, y) \wedge \neg \exists y R(y, x)) \right), \\ \psi(x) &:= \neg \exists y F(x, y) \wedge \bigwedge_{k \in ([D]^2)^2} L_k(x, x) \wedge \\ &\quad \forall y \left( y \neq x \rightarrow \bigwedge_{k \in ([D]^2)^2} \neg L_k(x, y) \wedge \bigwedge_{k \in ([D]^2)^2} \neg L_k(y, x) \right) \text{ and} \\ \chi(x) &:= \neg \exists y \bigvee_{k \in ([D]^2)^2} (L_k(x, y) \vee L_k(y, x)) \wedge \\ &\quad \bigwedge_{k \in ([D]^2)^2} \exists y_k \left( x \neq y_k \wedge F_k(x, y_k) \wedge \left( \bigwedge_{\substack{k' \in ([D]^2)^2 \\ k' \neq k}} \neg F_{k'}(x, y_k) \right) \wedge \forall y (y \neq y_k \rightarrow \neg F_k(x, y)) \right). \end{aligned}$$

Then  $\tilde{\varphi}_{\mathbb{Z}} := \varphi \wedge \forall x (\psi(x) \vee \chi(x)) \wedge \varphi_{\text{rotationMap}} \wedge \varphi_{\text{base}} \wedge \varphi_{\text{recursion}}$ . Hence it is sufficient to prove that  $A \cup A' \models \varphi$ ,  $A \cup A' \models \forall x (\psi(x) \vee \chi(x))$ ,  $A \cup A' \models \varphi_{\text{rotationMap}}$ ,  $A \cup A' \models \varphi_{\text{base}}$  and  $A \cup A' \models \varphi_{\text{recursion}}$ .

We first argue that  $A \cup A' \models \varphi$ . Let  $a \in U(A \cup A')$  be arbitrary and assume without loss of generality that  $a \in U(A)$ . Assume that  $A \cup A' \not\models \varphi_{\text{root}}(a) \wedge R(a, a)$ . Since

$\varphi_{\text{root}}(x) := \forall y \neg F(y, x)$  this implies that  $A \not\models \varphi_{\text{root}}(a) \wedge R(a, a)$ . Since  $A \models \varphi$  we get that  $A \models \exists^1 y F(y, a) \wedge \neg \exists y R(a, y) \wedge \neg \exists y R(y, a)$ . Hence there is an element  $b \in U(A)$  such that  $(b, a) \in F(A)$ . Furthermore, for every  $b' \in U(A)$ ,  $b' \neq b$  we have  $(b', a) \notin F(A)$ ,  $(a, b') \notin R(A)$  and  $(b', a) \notin R(A)$ . But because  $a$  cannot be in a tuple with any element in  $U(A')$  we get that  $A \cup A' \models \exists^1 y F(y, a) \wedge \neg \exists y R(a, y) \wedge \neg \exists y R(y, a)$ . Hence  $A \cup A' \models \varphi$ .

Next we prove that  $A \cup A' \models \forall x(\psi(x) \vee \chi(x))$ . Let  $a \in U(A \cup A')$  be arbitrary and assume without loss of generality that  $a \in U(A)$ . First assume that  $(a, b) \notin F(A \cup A')$  for every  $b \in U(A \cup A')$ . Since  $A$  is a substructure of  $A \cup A'$  this means that  $A \models \neg \exists y F(a, y)$ . But then  $A \not\models \bigwedge_{k \in ([D]^2)^2} \exists y_k (a \neq y_k \wedge F_k(a, y_k))$  which implies  $A \not\models \chi(a)$ . Since  $A \models \forall x(\psi(x) \vee \chi(x))$  this implies that  $A \models \psi(a)$ . Hence for every  $k \in ([D]^2)^2$  we have  $(a, a) \in L_k(A)$  and for every  $b \in U(A)$ ,  $b \neq a$  we have  $(a, b), (b, a) \notin F_k(A)$ . Since there are no tuples containing both elements from  $A$  and  $A'$  this directly implies that  $A \cup A' \models \psi(a)$ .

On the other hand, assume that there is  $b \in U(A \cup A')$  such that  $(a, b) \in F(A \cup A')$ . Since we are considering the disjoint union of  $A$  and  $A'$  this implies that  $b$  must be an element from  $A$ . Hence  $A \not\models \psi(a)$ . Since  $A \models \forall x(\psi(x) \vee \chi(x))$  this implies that  $A \models \chi(a)$ . Then for every  $k \in ([D]^2)^2$  there is an element  $b \in U(A)$  such that  $(a, b) \in F_k(A)$ ,  $(a, b) \notin F_{k'}(A)$  for every  $k' \in ([D]^2)^2$ ,  $k' \neq k$  and  $(a, b') \notin F_k(A)$  for every  $b' \in U(A)$ ,  $b' \neq b$ . But since in  $A \cup A'$  there are no tuples containing both elements from  $A$  and  $A'$  this implies that  $A \cup A' \models \chi(a)$ . In conclusion we proved that  $A \cup A' \models \forall x(\psi(x) \vee \chi(x))$ .

We now prove  $A \cup A' \models \varphi_{\text{rotationMap}}$ . Hence assume  $a, b \in U(A \cup A')$  are arbitrary elements. First consider the case that  $a, b$  are either both from  $U(A)$  or both from  $U(A')$ . Then if for some  $i, j \in [D]^2$  we have that  $(a, b) \in E_{i,j}(A \cup A')$  then  $(b, a) \in E_{j,i}(A \cup A')$  because  $A \models \varphi_{\text{rotationMap}}$  and  $A' \models \varphi_{\text{rotationMap}}$ . Now consider the case that  $|\{a, b\} \cap U(A)| = 1$ . Then  $(a, b) \notin E_{i,j}(A \cup A')$  and  $(b, a) \notin E_{j,i}(A \cup A')$  and hence  $A \cup A' \models \bigwedge_{i,j \in [D]^2} (E_{i,j}(a, b) \rightarrow E_{j,i}(b, a))$ . Therefore  $A \cup A' \models \forall x \forall y \left( \bigwedge_{i,j \in [D]^2} (E_{i,j}(x, y) \rightarrow E_{j,i}(y, x)) \right)$ .

Now consider an arbitrary element  $a \in U(A \cup A')$  and any  $i \in [D]^2$ . Without loss of generality assume  $a \in U(A)$ . Since  $A \models \varphi_{\text{rotationMap}}$  there must be an index  $j \in [D]^2$  and an element  $b \in U(A)$  such that  $(a, b) \in E_{i,j}(A)$ . Furthermore, for every  $b' \in U(A)$ ,  $b' \neq b$  we have  $(a, b') \notin E_{i,j}(A)$  and for every  $j' \in [D]^2$ ,  $j' \neq j$  and every  $\tilde{b} \in U(A)$  we have  $(a, \tilde{b}) \notin E_{i,j'}(A)$ . But since  $a \in U(A)$  it also holds that  $(a, b') \notin E_{i,j'}(A)$  for every  $b' \in U(A')$  and every  $j' \in [D]^2$ . Hence  $A \cup A' \models \bigvee_{j \in [D]^2} (\exists^1 y E_{i,j}(a, y) \wedge \bigwedge_{j' \in [D]^2, j' \neq j} \neg \exists y E_{i,j'}(a, y))$ . This concludes the proof of  $A \cup A' \models \varphi_{\text{rotationMap}}$ .

We now prove  $A \cup A' \models \varphi_{\text{base}}$ . Assume  $a \in U(A \cup A')$  is an arbitrary element such that  $A \cup A' \models \varphi_{\text{root}}(a)$ . Without loss of generality assume  $a \in U(A)$ . Since  $\varphi_{\text{root}}(x) := \forall y \neg F(y, x)$  and  $A \cup A' \models \varphi_{\text{root}}(a)$  we get that  $A \models \varphi_{\text{root}}(a)$ . Since  $A \models \varphi_{\text{base}}$  this means that for every  $i, j \in [D]^2$  we have  $(a, a) \in E_{i,j}(A)$  and  $(a, b), (b, a) \notin E_{i,j}(A)$  for every  $b \in U(A)$ ,  $b \neq a$ . Since further  $(a, b), (b, a) \notin E_{i,j}(A \cup A')$  for every  $b \in U(A')$  this implies that  $A \cup A' \models \bigwedge_{i,j \in [D]^2} \left( E_{i,j}(a, a) \wedge \forall y \left( a \neq y \rightarrow (\neg E_{i,j}(a, y) \wedge \neg E_{i,j}(y, a)) \right) \right)$ . Furthermore, since  $A \models \varphi_{\text{base}}$  and  $A \models \varphi_{\text{root}}(a)$  for every  $k, k' \in ([D]^2)^2$ ,  $i, i' \in [D]^2$  for which  $\text{ROT}_{H^2}(k, i) = (k', i')$  there are  $b, b' \in U(A)$  such that  $(a, b) \in F_k(A)$ ,  $(a, b') \in F_{k'}(A)$  and  $(b, b') \in E_{i,i'}(A)$ . Since  $A$  is a substructure of  $A \cup A'$  this proves that  $A \cup A' \models \varphi_{\text{base}}$ .

Finally we prove  $A \cup A' \models \varphi_{\text{recursion}}$ . Hence assume  $a, c \in U(A \cup A')$  are arbitrary elements.

Assume  $A \cup A' \models \neg \exists y F(a, y) \wedge \neg \exists y F(c, y)$  and assume without loss of generality that there is  $\tilde{a} \in U(A \cup A')$  such that  $(a, \tilde{a}) \in F(A \cup A')$ . Since there are no tuples containing both elements from  $A$  and  $A'$  we get that  $a, \tilde{a}$  are from the same structure. Without loss of generality assume  $a, \tilde{a} \in U(A)$ . Assume that for indices  $k'_1, k'_2 \in [D]^2, \ell'_1, \ell'_2 \in [D]^2$  and some element  $b \in U(A \cup A')$  we have  $(a, b) \in E_{k'_1, \ell'_1}(A \cup A')$  and  $(b, c) \in E_{k'_2, \ell'_2}(A \cup A')$ . As  $b$  also has to be in  $U(A)$  and  $A \models \varphi_{\text{recursion}}$  this implies that for every  $i, j, i', j' \in [D], k, \ell \in ([D]^2)^2$  for which  $\text{ROT}_H(k, i) = ((k'_1, k'_2), i')$  and  $\text{ROT}_H((\ell'_2, \ell'_1), j) = (\ell, j')$  there are elements  $a', c' \in U(A \cup A')$  such that  $(a, a') \in F_k(A \cup A'), (c, c') \in F_\ell(A \cup A')$  and  $(a', c') \in E_{(i, j), (j', i')}(A \cup A')$ . Hence  $A \cup A' \models \varphi_{\text{recursion}}$ .  $\triangleleft$

$\triangleright$  **Claim 30.** Every structure  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$  satisfies  $\varphi'_{\text{tree}}$ .

**Proof.** Let  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$ . Then there is  $k \in \{1, \dots, m\}$  such that  $A \in \mathcal{P}_{\rho_k}$ .

By definition,  $\varphi'_{\text{tree}} := \exists^{\leq 1} x \varphi_{\text{root}}(x) \wedge \varphi \wedge \forall x (\psi(x) \vee \chi(x))$ , where

$$\begin{aligned} \varphi &:= \forall x \left( (\varphi_{\text{root}}(x) \wedge R(x, x)) \vee (\exists^=1 y F(y, x) \wedge \neg \exists y R(x, y) \wedge \neg \exists y R(y, x)) \right), \\ \psi(x) &:= \neg \exists y F(x, y) \wedge \bigwedge_{k \in ([D]^2)^2} L_k(x, x) \wedge \\ &\quad \forall y \left( y \neq x \rightarrow \bigwedge_{k \in ([D]^2)^2} \neg L_k(x, y) \wedge \bigwedge_{k \in ([D]^2)^2} \neg L_k(y, x) \right) \text{ and} \\ \chi(x) &:= \neg \exists y \bigvee_{k \in ([D]^2)^2} (L_k(x, y) \vee L_k(y, x)) \wedge \\ &\quad \bigwedge_{k \in ([D]^2)^2} \exists y_k \left( x \neq y_k \wedge F_k(x, y_k) \wedge \left( \bigwedge_{\substack{k' \in ([D]^2)^2 \\ k' \neq k}} \neg F_{k'}(x, y_k) \right) \wedge \forall y (y \neq y_k \rightarrow \neg F_k(x, y)) \right). \end{aligned}$$

Thus, it is sufficient to prove that  $A \models \exists^{\leq 1} x \varphi_{\text{root}}(x)$ ,  $A \models \varphi$  and  $A \models \forall x (\psi(x) \vee \chi(x))$ .

To prove  $A \models \exists^{\leq 1} x \varphi_{\text{root}}(x)$  we note that by construction of  $\rho_k$  we have  $A \models \varphi_{\text{root}}(a)$  for any  $a \in U(A)$  for which  $(\mathcal{N}_2^A(a), a) \notin \tau_{d,2,\sigma}^k$ . Since  $\rho_k$  restricts the number of occurrences of elements of neighbourhood type  $\tau_{d,2,\sigma}^k$  to at most one, this proves that there is at most one  $a \in U(A)$  with  $A \models \varphi_{\text{root}}(a)$  and hence  $A \models \exists^{\leq 1} x \varphi_{\text{root}}(x)$ .

To prove  $A \models \varphi$ , let  $a \in U(A)$  be an arbitrary element. Since  $A \in \mathcal{P}_{\rho_k}$ , there is an  $i \in I_k$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^i$ . But then by definition, there exist  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Assume  $f$  is an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . First consider the case that  $A \models \varphi_{\text{root}}(a) := \forall y \neg F(y, a)$ . Assume there is  $\tilde{b} \in U(\tilde{A})$  such that  $(\tilde{b}, \tilde{a}) \in F(\tilde{A})$ . Since  $\tilde{b} \in \mathcal{N}_2^{\tilde{A}}(\tilde{a})$ , there must be an element  $b \in \mathcal{N}_2^A(a)$  such that  $f(b) = \tilde{b}$ . Since  $f$  is an isomorphism mapping  $a$  to  $\tilde{a}$ , this implies  $(b, a) \in F(A)$ , which contradicts  $A \models \varphi_{\text{root}}(a)$ . Hence  $\tilde{A} \models \varphi_{\text{root}}(\tilde{a})$ . Since  $\tilde{A} \models \varphi'_{\text{tree}}$ , it holds that  $\tilde{A} \models \varphi$ , which means that  $(\tilde{a}, \tilde{a}) \in R(\tilde{A})$ . But since  $f$  is an isomorphism mapping  $a$  onto  $\tilde{a}$ , this implies  $(a, a) \in R(A)$ . Now consider the case that  $A \not\models \varphi_{\text{root}}(a)$ . Then there is  $b \in U(A)$  with  $(b, a) \in F(A)$ . Since  $f$  is an isomorphism, this implies  $(f(b), \tilde{a}) \in F(\tilde{A})$ . Hence  $\tilde{A} \models \exists^=1 y F(y, \tilde{a}) \wedge \neg \exists y R(\tilde{a}, y) \wedge \neg \exists y R(y, \tilde{a})$ , as  $\tilde{A} \models \varphi$ . Now assume that there is  $b' \neq b$  such that  $(b', a) \in F(A)$ . Then  $f(b) \neq f(b')$  and  $(f(b), \tilde{a}), (f(b'), \tilde{a}) \in F(\tilde{A})$ . Since this contradicts  $\tilde{A} \models \exists^=1 y F(y, \tilde{a})$  we have  $A \models \exists^=1 y F(y, a)$ . Furthermore, assume that there is  $b' \in U(A)$  such that either  $(a, b') \in R(A)$  or  $(b', a) \in R(A)$ . Then either  $(\tilde{a}, f(b')) \in R(\tilde{A})$  or  $(f(b'), \tilde{a}) \in R(\tilde{A})$ , which contradicts  $\tilde{A} \models \neg \exists R(\tilde{a}, y) \wedge \neg \exists y R(y, \tilde{a})$ . Therefore  $A \models \neg \exists R(a, y) \wedge \neg \exists y R(y, a)$  which completes the proof of  $A \models \varphi$ .

We prove  $A \models \forall x (\psi(x) \vee \chi(x))$  by considering the two cases  $A \models \neg \exists y F(a, y)$  and  $A \models \exists y F(a, y)$  for each element  $a \in U(A)$ . For this, let  $a \in U(A)$  be any element. By

the construction of  $\rho_k$  there is  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Let  $f$  be an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . First consider the case that  $A \models \neg \exists y F(a, y)$ . If there was  $\tilde{b} \in U(\tilde{A})$  with  $(\tilde{a}, \tilde{b}) \in F(\tilde{A})$  then  $(a, f^{-1}(\tilde{b})) \in F(A)$  contradicting our assumption. Hence  $\tilde{A} \models \neg \exists y F(\tilde{a}, y)$  which implies that  $\tilde{A} \not\models \chi(\tilde{a})$ . But since  $\tilde{A} \models \varphi_{\mathbb{Z}}$ , it holds that  $\tilde{A} \models \forall x (\psi(x) \vee \chi(x))$ , which implies that  $\tilde{A} \models \psi(\tilde{a})$ . Hence  $(\tilde{a}, \tilde{a}) \in L_k(\tilde{A})$  for every  $k \in ([D]^2)^2$ . Since  $f$  is an isomorphism and  $f(a) = \tilde{a}$ , it holds that  $(a, a) \in L_k(A)$  for every  $k \in ([D]^2)^2$ , and hence  $A \models \bigwedge_{k \in ([D]^2)^2} L_k(a, a)$ . Furthermore, assume that there is  $b \in U(A)$ ,  $b \neq a$  and  $k \in ([D]^2)^2$  such that either  $(a, b) \in L_k(A)$  or  $(b, a) \in L_k(A)$ . Since  $f$  is an isomorphism this implies that either  $(\tilde{a}, f(b)) \in L_k(\tilde{A})$  or  $(f(b), \tilde{a}) \in L_k(\tilde{A})$  which contradicts  $\tilde{A} \models \chi(\tilde{a})$ . Hence  $A \models \forall y (y \neq a \rightarrow \bigwedge_{k \in ([D]^2)^2} \neg L_k(a, y) \wedge \bigwedge_{k \in ([D]^2)^2} \neg L_k(y, a))$  proving that  $A \models \psi(a)$ .

Now consider the case that there is an element  $b \in U(A)$  such that  $(a, b) \in F(A)$ . Since this implies that  $(\tilde{a}, f(b)) \in F(\tilde{A})$ , we get that  $\tilde{A} \not\models \psi(\tilde{a})$ , and hence  $\tilde{A} \models \chi(\tilde{a})$ . Now assume that there is  $b \in U(A)$  and  $k \in ([D]^2)^2$  such that either  $(a, b) \in L_k(A)$  or  $(b, a) \in L_k(A)$ . But then either  $(\tilde{a}, f(b)) \in L_k(\tilde{A})$  or  $(f(b), \tilde{a}) \in L_k(\tilde{A})$ , which contradicts  $\tilde{A} \models \chi(\tilde{a})$ . Hence  $A \models \neg \exists y \bigvee_{k \in ([D]^2)^2} (L_k(a, y) \vee L_k(y, a))$ . For each  $k \in ([D]^2)^2$ , let  $\tilde{b}_k \in U(\tilde{A})$  be an element such that  $\tilde{A} \models \tilde{a} \neq \tilde{b}_k \wedge F_k(\tilde{a}, \tilde{b}_k) \wedge (\bigwedge_{k' \in ([D]^2)^2, k' \neq k} \neg F_{k'}(\tilde{a}, \tilde{b}_k)) \wedge \forall y (y \neq \tilde{b}_k \rightarrow \neg F_k(\tilde{a}, y))$ . Since  $f$  is an isomorphism, this implies that  $a \neq b_k := f^{-1}(\tilde{b}_k)$ ,  $(a, b_k) \in F_k(A)$  and  $(a, b_k) \notin F_{k'}(A)$ , for each  $k' \in ([D]^2)^2, k' \neq k$ . Furthermore, assume there is  $b \in U(A)$ ,  $b \neq b_k$  such that  $(a, b) \in F_k(A)$ . Since  $f$  is an isomorphism, this implies  $f(b) \neq f(b_k) = \tilde{b}_k$  and  $(\tilde{a}, \tilde{b}) \in F_k(\tilde{A})$ , which contradicts  $\tilde{A} \models \forall y (y \neq \tilde{b}_k \rightarrow \neg F_k(\tilde{a}, y))$ . Hence  $A \models \forall y (y \neq b_k \rightarrow \neg F_k(a, y))$  and therefore concluding that  $A \models \chi(a)$ . This proves that in either case  $A \models \psi(a) \vee \chi(a)$  and therefore  $A \models \forall x (\psi(x) \vee \chi(x))$ .  $\triangleleft$

▷ **Claim 31.** Every structure  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$  satisfies  $\varphi_{\text{rotationMap}}$ .

**Proof.** Let  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$ . Then there is a  $k \in \{1, \dots, m\}$  such that  $A \in \mathcal{P}_{\rho_k}$ .

By definition,  $\varphi_{\text{rotationMap}} = \varphi \wedge \psi$ , where

$$\begin{aligned} \varphi &:= \forall x \forall y \left( \bigwedge_{i, j \in [D]^2} (E_{i, j}(x, y) \rightarrow E_{j, i}(y, x)) \right) \text{ and} \\ \psi &:= \forall x \left( \bigwedge_{i \in [D]^2} \left( \bigvee_{j \in [D]^2} (\exists^=1 y E_{i, j}(x, y) \wedge \bigwedge_{\substack{j' \in [D]^2 \\ j' \neq j}} \neg \exists y E_{i, j'}(x, y)) \right) \right). \end{aligned}$$

Thus, it is sufficient to prove that  $A \models \varphi$  and  $A \models \psi$ .

To prove  $A \models \varphi$ , assume towards a contradiction that there are  $a, b \in U(A)$  such that for some pair  $i, j \in [D]^2$ , we have that  $(a, b) \in E_{i, j}(A)$ , but  $(b, a) \notin E_{j, i}(A)$ . By construction of  $\mathcal{P}_{\rho_k}$ , there is a structure  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Assume  $f$  is an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Note that  $f(b)$  is defined since  $b$  is in the 2-neighbourhood of  $a$ . Furthermore since  $f$  is an isomorphism,  $(a, b) \in E_{i, j}(A)$  implies  $(\tilde{a}, f(b)) \in E_{i, j}(\tilde{A})$ , and  $(b, a) \notin E_{j, i}(A)$  implies  $(f(b), \tilde{a}) \notin E_{j, i}(\tilde{A})$ . Hence  $\tilde{A} \not\models \varphi$ , which contradicts  $\tilde{A} \models \varphi_{\text{rotationMap}}$ .

To prove  $A \models \psi$ , assume towards a contradiction that there is an  $a \in U(A)$  and  $i \in [D]^2$  such that  $A \not\models \exists^=1 y E_{i, j}(a, y) \wedge \bigwedge_{\substack{j' \in [D]^2 \\ j' \neq j}} \neg \exists y E_{i, j'}(a, y)$  for every  $j \in [D]^2$ . We know that there is a structure  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Let  $f$  be an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Since  $\tilde{A} \models \psi$ , there must be  $j \in [D]^2$

such that  $\tilde{A} \models \exists^{=1} y E_{i,j}(\tilde{a}, y) \wedge \bigwedge_{j' \in [D]^2, j' \neq j} \neg \exists y E_{i,j'}(\tilde{a}, y)$ . Hence there must be  $\tilde{b} \in U(\tilde{A})$  such that  $(\tilde{a}, \tilde{b}) \in E_{i,j}(\tilde{A})$ , which implies that  $(a, f^{-1}(\tilde{b})) \in E_{i,j}(A)$ . Since we assumed that  $A \not\models \exists^{=1} y E_{i,j}(a, y) \wedge \bigwedge_{j' \in [D]^2, j' \neq j} \neg \exists y E_{i,j'}(a, y)$ , there must be either  $b \neq f^{-1}(\tilde{b})$  with  $(a, b) \in E_{i,j}(A)$ , or there must be  $j' \in [D]^2, j' \neq j$  and  $b' \in U(A)$  such that  $(a, b') \in E_{i,j'}(A)$ . In the first case  $(\tilde{a}, f(b)) \in E_{i,j}(\tilde{A})$ , since  $f$  is an isomorphism. But then  $\tilde{A} \models \exists^{=1} y E_{i,j}(\tilde{a}, y)$ , which is a contradiction. In the second case, we get that  $(\tilde{a}, f(b')) \in E_{i,j'}(\tilde{A})$ . But then  $\tilde{A} \models \bigwedge_{j' \in [D]^2, j' \neq j} \neg \exists y E_{i,j'}(\tilde{a}, y)$ , which is a contradiction. Hence  $A \models \varphi \wedge \psi$ .  $\triangleleft$

▷ **Claim 32.** Every structure  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$  satisfies  $\varphi_{\text{base}}$ .

**Proof.** Let  $A \in \bigcup_{1 \leq k \leq m} \mathcal{P}_{\rho_k} \setminus \{A_\emptyset\}$ . Then there is a  $k \in \{1, \dots, m\}$  such that  $A \in \mathcal{P}_{\rho_k}$ .

By definition,  $\varphi_{\text{base}} := \forall x (\varphi_{\text{root}}(x) \rightarrow (\varphi(x) \wedge \psi(x)))$ , where

$$\begin{aligned} \varphi(x) &:= \bigwedge_{i,j \in [D]^2} \left( E_{i,j}(x, x) \wedge \forall y \left( x \neq y \rightarrow (\neg E_{i,j}(x, y) \wedge \neg E_{i,j}(y, x)) \right) \right) \text{ and} \\ \psi(x) &:= \bigwedge_{\substack{\text{ROT}_{H^2}(k,i)=(k',i') \\ k,k' \in ([D]^2)^2 \\ i,i' \in [D]^2}} \exists y \exists y' (F_k(x, y) \wedge F_{k'}(x, y') \wedge E_{i,i'}(y, y')). \end{aligned}$$

Thus, it is sufficient to prove that  $A \models \varphi(a)$  and  $A \models \psi(a)$  for every  $a \in U(A)$  for which  $A \models \varphi_{\text{root}}(a)$ . Therefore assume  $a \in U(A)$  is any element such that  $A \models \varphi_{\text{root}}(a)$ . Because  $A \in \mathcal{P}_{\rho_k}$  there is an  $i \in I_k$  such that  $(\mathcal{N}_2^A(a), a) \in \tau_{d,2,\sigma}^i$ . Then by definition there is a structure  $\tilde{A} \models \varphi_{\mathbb{Z}}$  and  $\tilde{a} \in U(\tilde{A})$  such that  $(\mathcal{N}_2^A(a), a) \cong (\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Let  $f$  be an isomorphism from  $(\mathcal{N}_2^A(a), a)$  to  $(\mathcal{N}_2^{\tilde{A}}(\tilde{a}), \tilde{a})$ . Assume that there is an element  $\tilde{b} \in U(\tilde{A})$  such that  $(\tilde{b}, \tilde{a}) \in F(\tilde{A})$ . Since  $f$  is an isomorphism and  $\tilde{b} \in \mathcal{N}_2^{\tilde{A}}(\tilde{a})$  we get that  $(f^{-1}(\tilde{b}), a) \in F(A)$  which contradicts that  $A \models \varphi_{\text{root}}(a)$  as  $\varphi_{\text{root}}(x) := \forall y \neg F(y, x)$ . Hence there is no element  $\tilde{b} \in U(\tilde{A})$  such that  $(\tilde{b}, \tilde{a}) \in F(\tilde{A})$  which implies that  $\tilde{A} \models \varphi_{\text{root}}(\tilde{a})$ . But since  $\tilde{A} \models \varphi_{\mathbb{Z}}$  we have that  $\tilde{A} \models \varphi_{\text{base}}$  and hence  $\tilde{A} \models \varphi(\tilde{a})$  and  $\tilde{A} \models \psi(\tilde{a})$ .

To prove  $A \models \varphi(a)$  first observe that  $(a, a) \in E_{i,j}(A)$  for every  $i, j \in [D]^2$  since  $\tilde{A} \models \varphi(\tilde{a})$  implies that  $(\tilde{a}, \tilde{a}) \in E_{i,j}(\tilde{A})$  for every  $i, j \in [D]^2$  and  $f$  is an isomorphism mapping  $a$  onto  $\tilde{a}$ . Assume that there is an element  $b \in U(A)$ ,  $b \neq a$  and indices  $i, j \in [D]^2$  such that either  $(a, b) \in E_{i,j}(A)$  or  $(b, a) \in E_{i,j}(A)$ . Since  $b \in \mathcal{N}_2^A(a)$  and  $f$  is an isomorphism we get that  $f(b) \neq f(a) = \tilde{a}$  and either  $(\tilde{a}, f(b)) \in E_{i,j}(\tilde{A})$  or  $(f(b), \tilde{a}) \in E_{i,j}(\tilde{A})$ . But this contradicts  $\tilde{A} \models \varphi(\tilde{a})$  and hence  $A \models \varphi(a)$ .

We now prove  $A \models \psi(a)$ . Let  $k, k' \in ([D]^2)^2$  and  $i, i' \in [D]^2$  such that  $\text{ROT}_{H^2}(k, i) = (k', i')$ . Since  $\tilde{A} \models \psi(\tilde{a})$  there must be elements  $\tilde{b}, \tilde{b}' \in U(\tilde{A})$  such that  $(\tilde{a}, \tilde{b}) \in F_k(\tilde{A})$ ,  $(\tilde{a}, \tilde{b}') \in F_{k'}(\tilde{A})$  and  $(\tilde{b}, \tilde{b}') \in E_{i,i'}(\tilde{A})$ . But since  $\tilde{b}, \tilde{b}' \in \mathcal{N}_2^{\tilde{A}}(\tilde{a})$  we get that  $f^{-1}(\tilde{b})$  and  $f^{-1}(\tilde{b}')$  are defined and since  $f$  is an isomorphism we get that  $(a, f^{-1}(\tilde{b})) \in F_k(A)$ ,  $(a, f^{-1}(\tilde{b}')) \in F_{k'}(A)$  and  $(f^{-1}(\tilde{b}), f^{-1}(\tilde{b}')) \in E_{i,i'}(A)$ . Hence  $A \models \exists y \exists y' (F_k(a, y) \wedge F_{k'}(a, y') \wedge E_{i,i'}(y, y'))$  for any  $k, k' \in ([D]^2)^2$  and  $i, i' \in [D]^2$  such that  $\text{ROT}_{H^2}(k, i) = (k', i')$  which implies that  $A \models \psi(a)$ .  $\triangleleft$